OULU BUSINESS SCHOOL

UNIVERSITY OF OULU

Jaume Gomà Sala

FUNDAMENTALS AND DYNAMICS OF HOUSE PRICES: THE CASE OF THE CAPITAL REGION OF HELSINKI (1995-2019)

Master’s Thesis

Department of Economics, Accounting and Finance

May 2021

UNIVERSITY OF OULU ABSTRACT OF THE MASTER'S THESIS

Oulu Business School

Unit			
Department of Economics, Accounting and Finance			
Author Gomà Sala Jaume		Supervisor Korhonen, M. Professor	
Title Fundamentals and dynamics of house prices: the case of the Capital Region of Helsinki (1995-2019)			
Subject Economics	Type of the degree Master's Thesis	Time of publication May 2021	Number of pages 93 + 6 pcs

Abstract

This thesis attempts to answer whether house prices are at their long-term equilibrium level, whether house prices converge towards their long-term equilibrium level and what drives house prices in the Capital Region of Helsinki (CRH). Our sample period covers 1995Q1-2019Q4 on behalf of the stability of the estimated coefficients, avoiding the distortion from periods of great volatility and structural changes (financial deregulation during the 1980s and the economic crisis of the 1990s).

This analysis is sustained on Oikarinen's (2009) life-cycle model, although keeping an eye on other relevant research (e.g., Meen 1990, Hort 1998, Meen 2002 and Oikarinen 2007). We run Johansen cointegration test to test for cointegration relationships and thereafter we estimate three different Vector Error Correction Models. Thus simple, these models capture the fundamentals underlying house prices in the CRH. The demand side of the housing market is catered by the real rental price determinants (such as aggregate income and some demographic variables), the real user cost and a loan stock variable. Construction costs and housing stock (real rental price determinant) cater for the supply side of the market.

Demand side fundamentals, aggregate income (and therefore, demographic factors), loan-to-GDP ratio, lagged house prices and real user costs are driving house price in the Capital Region of Helsinki. Supply side fundamentals are proven irrelevant as house price determinants. Based on our "best-fitted" model, results suggest the existence of a long-term relationship among house prices, aggregate income (and therefore demographic factors) and loan-to-GDP ratio. Whereas in the short-term, only lagged values of house prices and real user cost show short-term causality on house prices.

We find that lagged values of house prices are far from following a random walk process. As DiPasquale and Wheaton (1994) state. The three models show short-term causality between lagged values of house prices and current prices. Estimated income elasticities, range from 0.492 and 0.697 and loan-to-GDP ratio elasticities range within 0.237 and 0.302, in line with previous research in Finland (see e.g., Takala & Pere 1991, Barot & Takala 1998 or Oikarinen 2009).

Moreover, this thesis' results suggest that house prices converge towards the long-term equilibrium level. Real house prices converge, albeit at a slow pace, about 4.5%-6.8% a quarter. Results are in the line of previous research conducted in Finland (see e.g., Takala & Pere 1991, Barot & Takala 1998 or Oikarinen 2009). Hence, the assumption of market clearing in housing markets has to be relaxed as some authors suggest (see e.g., Case & Schiller 1990, Mankiw & Weil 1989). Some degree of inefficiency in housing market describes better why market prices are sluggish in CRH.

Keywords

Johansen Cointegration test, Granger-causality, Sluggishness, VECM.

Additional information

CONTENTS

1 INTRODUCTION

1.1 Background

High volatility in house prices is a phenomenon observed in many countries, very likely culminating in a housing market crisis. Furthermore, there is a credible threat that housing market crisis spreads to other sectors and ends in a long-lasting recession adversely affecting gross domestic product (GDP), employment and household’s wealth. (See e.g., Oikarinen 2005, Taipalus 2006, Steiner 2010 and Bourassa et al. 2019)

Multiple interdependencies between real estate markets and macroeconomics may harm domestic economies and eventually households. This way, real estate prices and their shifts are of great significance as they determine a great deal of household´s source of wealth and their consumption decisions (Poterba, Weil & Schiller, 1991) as they determine as well household indebtedness levels (Gimeno & Martínez-Carrascal, 2010). DiPasquale and Wheaton (1992) identify property as the largest component of household´s wealth and the largest ingredient of household’s investment in United States.

Not in vain, policy makers should be preoccupied about housing prices volatility:

> *Firstly, housing composes the majority of many households’ wealth, and the “wealth effect” of housing on consumption is significant (...). Hence, a decline in housing price level leads to less consumption. Secondly, a drop in housing prices is likely to have a negative effect on housing construction, and thus on aggregate output. In addition, a notable fall in housing prices would affect the banking sector by inducing unanticipated losses for mortgage lenders, which would strain the financial system* (Oikarinen 2005, p. 1).

In the recent years, Bourassa et al. (2019) recall how house price bubble in United States resulted into the US subprime crisis that spread worldwide as a major financial crisis, and how some economies exposed to price volatility have suffered long-lasting consequences such as Spain or Ireland.

Poterba et al. (1991) describe house markets as “inefficient” when compared to other asset markets, drawing special attention on house prices fluctuations, hence there are reasons to believe that prices are not a good sing of the market. Although they find evidence on its predictive power, describing prices as “forwards looking”, they recognize as well that prices are “containing information not found in other assets” (p. 174). In their opinion, it is not surprising that, “...several episodes of sharp increases and decreases in real house prices can be subject to speculative bubbles” (p. 145).

Therefore, an important question arises. How do we know whether house prices are consistent with market fundamentals or are a product of irrational behaviour? In words of Stiglitz (1991): “…if the reason that the price is high today is only because investors believe that the selling price will be high tomorrow-when "fundamental" factors do not seem to justify such a price-then a bubble exists” (p. 13).

More specifically, “a house price bubble occurs when homeowners have unreasonably high expectations about future capital gains, leading to perceive their user cost to be lower that actually is and thus pay “too much” to purchase a house today” (Himmelberg, Mayer and Sinai, 2005, p. 74).

It is this thesis ambition to understand those market fundamentals lying behind Capital Region of Helsinki (CRH) house prices whereas to unveil whether house prices are aligned with their long-term equilibrium levels. It is worth mentioning that might it be presence of mispricing in the CRH housing market it would be extremely unfortunate “to interpret marked price declines which occur without any apparent new information as the breaking of a bubble” (Stiglitz, 1991, p. 16).

1.2 The Finnish and Capital Region area housing market.

During 1980 and 1990 Finnish housing market undergoes major changes. They coincide with the economic upturn of late 1980s and the economic turmoil of Finnish banking crisis of 1990s. Especially in late 1980s, favoured by a positive international and domestic economic environment, Finland experienced of an era of financial deregulation resulting in the elimination of capital import controls and credit rationing and therefore the credit expands (Loikkanen and Lönnqvist, 2007). This gave access to households and construction firms to financing. Interest rates were kept at abnormally low levels, not limiting borrowing. Loikkanen and Lönnqvist (2007) say "For households, the combination of (de)regulation and the tax deductibility of interest payments resulted in negative real rates of interest" (p. 67).

By that time, Finnish housing market was highly regulated in one form or another[1], resulting in an excess of demand for rental and owner-occupied dwelling. In pursuit of equity, housing market prices were constrained by regulation. The access to credit led to a housing market bubble and a construction boom. As a result, house prices boosted.

The financial crisis of beginning of 1990s led to a devaluation of the Markka, rising interest rates and lending constraints. All these factors drove house prices to collapse. Households experienced unemployment and loss of wealth. Construction companies halted developments and retained unsold stocks, culminating into bankruptcies or merges. (Loikkanen and Lönnqvist, 2007)

Two notable events took place during the late stage of the crisis and its aftermath: the 1993 tax reform and the end of rent controls. Tax reform, from one standpoint, "limited the deductibility of interest payments of housing loans in taxation, narrowing the gap between real interest rate and after-tax interest rate significantly" (Loikkanen and Lönnqvist, 2007, p. 69) and secondly income from owner-occupied housing was eliminated. This in practice, according to Oikarinen (2007, p. 281) distorted the relationship among different asset prices since capital gains from owner-occupied homes are virtually tax-exempt, as to other capital gains are taxed under capital income tax. The second one, rent controls in the private market were progressively relaxed by 1993 by lifting rent controls from new rental agreements, culminating in 1995, when rent controls were totally abolished.

Loikkanen and Lönnqvist (2007) describe the property of the land in CRH. Around 66% of land area is city-owned in Helsinki. Contrary to Helsinki, in the rest of CRH (Espoo, Kauniainen and Vantaa) most of the land is privately owned: municipalities

[1] Both owner-occupying and rental housing.

owning less than a fourth of land. This determines price and availability by municipalities (leased land).

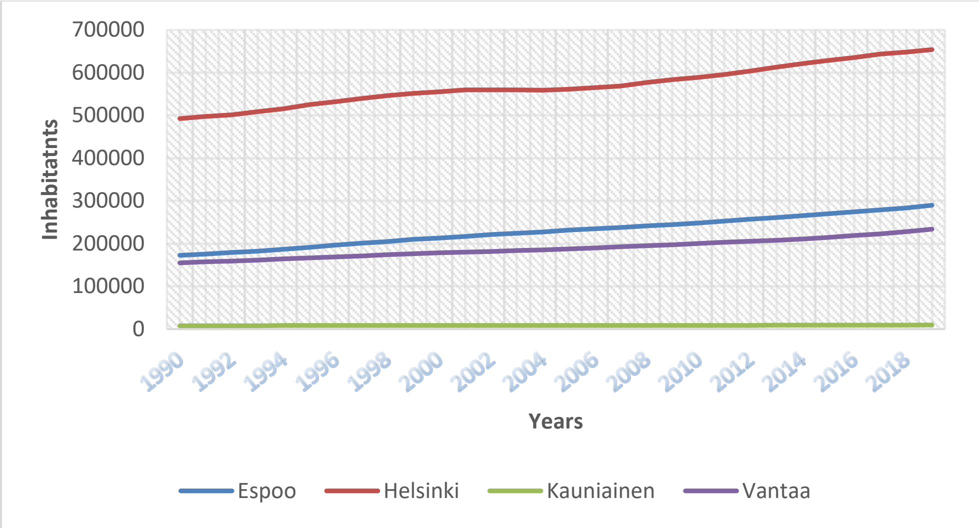

Figure 1. Evolution of population in CRH 1990-2019. (Source Tilastokeskus)

According to Tilastokeskus[2] (OSF), internal migration movements to urban areas, has been the tendency for decades, for instance in the beginning of 1990 about 63% of Finnish population lived in urban areas whereas by 2019 about 72%. Figure 1 depicts the population in CRH (number of inhabitants), the population of CRH has increased since 1990 above 40%. By cities, in Espoo has increased about 67%, in Vantaa doubled whereas in Helsinki or Kauniainen showed more moderated increase (33% and 24% respectively). Internal migrations have mostly settled in Helsinki area, Oulu, Jyvaskyla, Turku, and Tampere.

Since economic recovery (after 1995) and especially since Finland joined European Union (1995), external migration has become a new phenomenon, based on OSF statistics. Most of foreign population migrates to big cities, notably CRH, which concentrates most of investment and job opportunities. This way most of later

[2] Official Statistics of Finland (OSF): Migration [e-publication]. ISSN=1797-6782. Helsinki: Statistics Finland [referred: 16.3.2021].
Retrieved from https://www.stat.fi/til/muutl/2019/02/muutl_2019_02_20201221_tie_001_en.html

population gain in urban areas is sustained by immigrants, which coincides with a sustained increase in house prices for the last decade.

Another feature in the CRH is the tenure status, renting is becoming more common[3] in later years. In 2005 the number of the household units in the CRH living in owner-occupied dwellings were about 52.6%, in rented dwellings 42.4% of households whereas only 2.8% in right of occupancy dwellings. In 2019 the number of households living in owner-occupied dwellings was slightly inferior, 49% of households. Whereas those households living in rented dwellings was over 45%. The number of units living in right of occupancy dwelling rose up to 3.7%. By cities, in Espoo, Kauniainen and Vantaa most of households live in owner-occupied dwellings whereas in Helsinki the most common tenure status is rented dwelling (49.2%) followed by owner-occupied dwellings (45.2%).

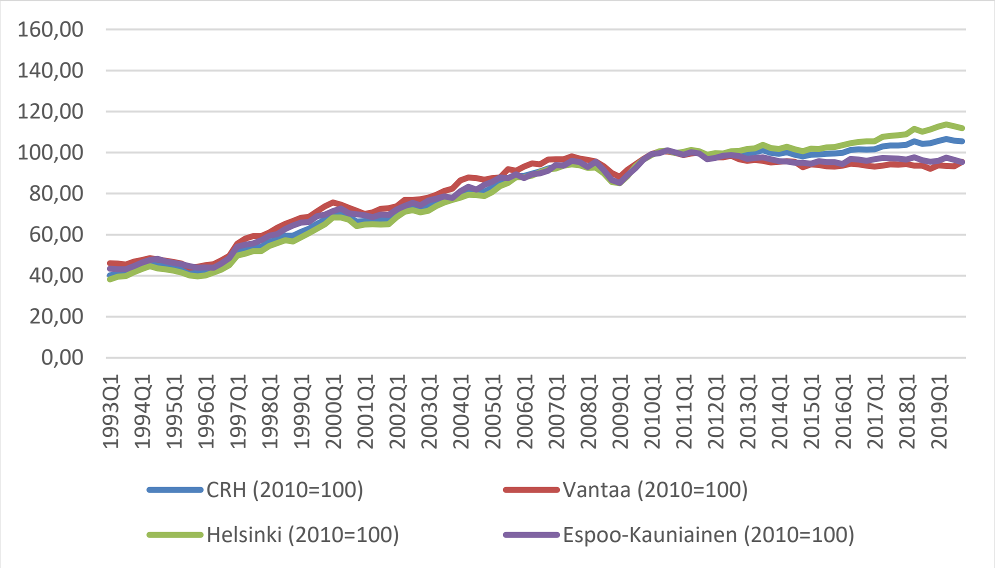

Figure 2. Real Price Indices 1993-2020. (Source Tilastokeskus)

As Figure 2 shows, from the beginning of 1996 until the second quarter of 2008 real house prices in CRH increased 81% (Helsinki 82%, Espoo-Kauniainen 80% and Vantaa 76%). From the second quarter of 2009 (2009Q2) until the end of 2019 (2019Q4), prices in CRH increased approximately 21%, while in Helsinki climbed

[3] Official Statistics of Finland (OSF): Dwellings and housing conditions [e-publication]. ISSN=1798-6761. Overview 2017. Helsinki: Statistics Finland [referred: 16.3.2021]. Retrieved from: http://www.stat.fi/til/asas/2017/01/asas_2017_01_2018-10-10_tie_002_en.html

28%. Real house prices in Espoo/Kauniainen (12%) show similar trend than Vantaa (8%), growing both cities at a slower pace than the CRH average.

Since the beginning of 1996 until the financial crisis of 2008, real house prices grew in Helsinki, Espoo/Kauniainen and Vantaa at a higher pace than Finnish average: what drove prices growing at such fast pace? Oikarinen (2005) provides some answers to this question. He finds real loan-to-GDP ratio, income, and income expectations as house price´s drivers. Although, he finds no traces of any house price bubble, he predicts that house prices will continue growing in the following years (as it happened). By contrast, especially after 2009Q2, real house prices in Helsinki have grown more than twice as fast than Finnish average (see Figure 2).

1.3 Motivation

Attempting to answer some of the following (and further) questions is what this research aims to do:

a) Are house prices in the CRH at their long-term equilibrium level?
b) Do house prices converge towards their long-term equilibrium level?
c) What drives house prices in the CRH?

Previous research has been done about the determinants of prices, with strong focus on CRH or Helsinki, and attempted to answer some of these questions already. The aim of this thesis is to throw some more light on the insights of this area and to contribute to research with its grain of sand.

In most of the recent research we had access to, the sample period ends around the first decade of the twenty first century: e.g., Oikarinen (2009) covers from 1975Q1 to 2006Q2, André and García (2012) cover 1995 to 2010, or more recently, Oikarinen (2014) comprises 1988Q1 to 2008Q2 or Bourassa, Hoesli & Oikarinen (2019) span from 1975Q1 to 2012Q4. To this purpose, this paper analyses some of the determinants of house pricing in the CRH from 1995Q1 until 2019Q4.

The span of time covered should be sufficient to conduct sound empirical research. We have assumed, in the same vein than André and García (2012), that our estimation begins in 1995Q1 on behalf of the stability of the estimated coefficients, avoiding the distortion from periods of great volatility and structural changes (financial deregulation during the 1980s and the economic crisis of the 1990s). It is not going beyond 2019Q4 since some of the available 2020 data released by Tilastokeskus (http://www.tilastokeskus.fi/) is preliminary and therefore not yet reliable. This lack of reliability suggest that the most prudent course is to not take year 2020 as a study year, as it is likely to introduce some bias and lead to wrong conclusions.

1.4 Prior research and contribution

Previous research on house prices in the CRH included in this thesis begins with Kuismanen, Laakso and Loikkanen (1999), Oikarinen (2005), Oikarinen (2007), Oikarinen (2009), Oikarinen (2014) and Bourassa et al. (2919). Other authors contributed to the research on house prices at a national level such as Takala and Pere (1991), Barot and Takala (1998), Oikarinen (2006), Taipalus (2006), André and García (2012) and Oikarinen (2012). We are aware of other authors who contributed on housing price research in Finland, such as Laakso (2000) or Kosonen (1997), but unfortunately, we have not gained access to their research, and although we are aware of their contribution, we avoided using secondary sources to report their findings.

Our econometrics analysis is sustained on Oikarinen's (2009) life-cycle model, although keeping an eye on other relevant research (e.g., Meen 1990, Hort 1998, Meen 2002 and Oikarinen 2007). We run Johansen cointegration test to test for cointegration relationships and thereafter we estimate three different Vector Error Correction Models (VECM). Thus simple, these models capture the fundamentals underlying house prices in the CRH: the long-term equilibria and short-term dynamics (speed of adjustment) and the house price fundamentals. The demand side of the housing market is catered by some real rental price determinants (such as aggregate income and some demographic variables), the real user cost and the loan stock variable. The construction costs and housing stock (real rental price determinant) cater for the supply side of the market.

Supply side fundamentals are proven irrelevant (other research comes to the same view i.e., Oikarinen 2009). We suspect that housing stock data may induce multicollinearity in our analyses. Although real construction costs are constant during the whole sample period, there is a rising trend of real house prices. This implies that there must be other reasons for the recognised trend.

Demand side fundamentals such as the aggregate income, and thus demographic factors, loan-to-GDP ratio, lagged values of house prices and of the real user cost are driving house price in the Capital Region of Helsinki. Based on our "best-fitted" model (Model 3), results suggest the existence of a long-term relationship among house price, aggregate income (and therefore demographic factors) and loan-to-GDP ratio. Whereas in the short-term, lagged value of house prices and lagged values of real user cost show short-term causality on house prices (supported by Granger-causality).

Estimated income elasticities, range from 0.492 and 0.697, and are in the same vein than previous research, both at national and international level (see e.g., Abraham & Hendershott 1996, Kuismanen et al. 1999, Meen & Wallace 2003, Oikarinen 2005 or Oikarinen 2009). We estimate a loan-to-GDP ratio elasticity within 0.237 and 0.302 in line with previous research (see e. g. Oikarinen 2005, 2009 and 2012).

The results from the preferred model (Model 3) suggest that house prices converge towards the long-term equilibrium level but at a slow pace, about 7% a quarter ($\alpha = -0.068$). Other estimated models show a speed of quarterly adjustment within 4.5% and 5.4%.

Results are in the line of previous research conducted in Finland (see e.g., Takala & Pere 1991, Barot & Takala 1998 or Oikarinen 2009). Hence, the assumption of market clearing in housing markets has to be relaxed as some authors suggest (see e.g., Case & Schiller 1990 or Mankiw & Weil 1989). Some degree of inefficiency in housing market describes better why market prices are sluggish in CRH.

Finally, the three models show short-term causality between current prices and their own lagged values. Hence, in the same vein as DiPasquale and Wheaton (1994), house prices are highly predictable based on positive serial price correlation.

1.5 Master's thesis description

This thesis begins by describing some theoretical framework based on DiPasquale and Wheaton (1992), DiPasquale and Wheaton (1994), Poterba (1984), Himmelberg et al. (2005) and Meen (1990). These papers could be considered seminal papers, since they are widely referenced and have served as baseline for later research. These papers have proven to be a sound and solid foundation for house price research. Following these frameworks, section covers a survey on relevant literature review, both from Finland and abroad.

Section 3 covers the technical aspects and insights of the used data and the methodology and the justification to support our research methods. The following section presents the results for CRH, discusses, and compares them with other relevant research. As Oikarinen (2005) states, it is reasonable to think that long-term equilibrium exists among housing prices and some fundamental variables. As it would be reasonable to study their short-term adjustment dynamics and understand what eventually drive them out of their long-term equilibrium, or why it takes so long to converge towards the long-term equilibrium levels (sluggishness of prices).

Finally, Section 5 launches a set of conclusions and different thoughts that originate from our research, considered interesting to share as well as some considerations about further research.

2 ANALYTIC FRAMEWORK

In this section begins with some introduction to the theoretical framework developed by DiPasquale and Wheaton (1992). According to DeSalvo (2017) the DiPasquale and Wheaton (1992) framework stands among the most popular macro-models for real estate to examine long-term equilibrium house prices. The following part presents the DiPasquale and Wheaton (1994), Poterba's (1984) asset-market model of the housing market, the user cost of housing from Himmelberg, Mayer and Sinai (2005) and the life-cycle approach of Meen (1990). Thereupon, this section ends with a literature review that covers some relevant empirical research on the field.

2.1 DiPasquale and Wheaton (1992): A stock-flow approach

DiPasquale and Wheaton model, is in words of Colwell (2002) "a model of long-run equilibrium in the aggregate real estate market with a connection to the financial capital market", therefore validating the existence of a long-term equilibrium price. Both DeSalvo (2017) and Colwell (2002) acknowledge its pedagogical power and concentration as a flexible stock-flow model, although e. g. Colwell criticises the lack of expectations in the model. A posteriori paper from DiPasquale and Wheaton (1994) settle some deficiencies with a more realistic perspective i.e., with the adoption of price formation expectations (incorporating the possibility of backward-looking expectations). DiPasquale and Wheaton model fail to describe into detail short-term dynamics of the market. This is later addressed by authors incorporating cointegration analysis, Error Correction Models (ECM) and VECM. Prior to DiPasquale and Wheaton (1992), a comprehensive collection of stock-flow models is to be found (see e.g., Smith 1969; Higgins 1972; Fischer 1992).

2.1.1 Analytic framework for real estate market

Real estate market is divided in two different interdependent markets: the market for real estate space, property market, and the market for real estate assets, the asset market. The first one comprises the market for property occupied by its owners. The latter comprises the market that engages in buying and selling activities as well as the rental market.

The goods offered within the market are durable capital goods. As in any other market, supply and demand determine the quantity available and the price levels. Therefore, prices depend negatively on its availability in the market (supply) and positively on the desire of possessing such an asset by households (demand).

The market supply depends on market prices with respect to construction costs (production costs). In this context, construction costs represent the costs of replacing or supplying new stock. In the short run, prices and construction costs may differ due to the rigidities from supply, but in the long run, market prices and construction costs coincide. If prices are temporally above (for instance due to increase in demand) or below construction costs, market forces take their course over prices (or construction costs) and as time goes by converge towards each other. Any gap between construction costs and housing prices triggers construction activity.

The market demand on property market originates from households willing to occupy space (set in some measure, in the original paper in squared feet). This way, household demand depends on their income levels and on the cost of occupying the space (rent) in proportion to the cost of consuming other goods. Hence, if households are tenants, rents are determined by rental agreements, by contrast if households are owners rent is measured through some annualised cost related to the possession of the asset.

In the property market, supply and demand for space determine rental level. Market supply is given by the asset market size, while demand depends on rents and other exogenous variables (number of households, permanent income or level of space). When demand increases, for instance due to an increase in the number of households, if supply remains the same, rents tend to rise and vice versa. Supply for space correlates negatively with rents, by expanding the amount of space, rents are pushed down.

Property market and asset market are linked through the rental level and the construction sector. Rents, once determined in the market for space, shape the demand for real estate assets. Hence, any variation in rents influence the expected flow of income of investors. Increases have a positive effect on the demand for real estate assets and vice versa. Through construction sector, changes in the level of supply of real estate assets, for instance by increasing the construction level, drives rents (real estate space market) and prices (real estate asset market) down.

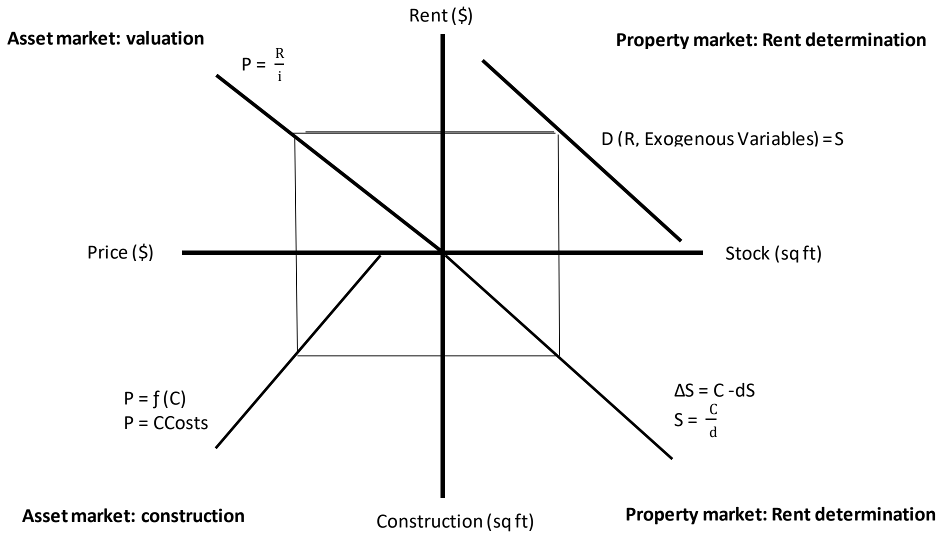

Figure 3. Four Quadrant model (adapted from DiPasquale and Wheaton (1992), p. 188)

All relationships discussed thus far, can be illustrated by depicting a four-quadrant diagram as in DiPasquale and Wheaton (1992) (Figures). This diagram describes the long-term equilibrium in the real estate space and real estate asset markets. Left-hand quadrants represent the market for real estate assets (North-West and South-West quadrants) and the right-hand quadrants constitute the market for real estate space (North-East and South-East quadrants).

The North-East quadrant determines the rental level. Its axes depict rents in the vertical axis and stock of space in the horizontal axis. The demand for space (D) in this quadrant equals the supply for space (S). In the equilibrium, rent (R) is determined. Demand is a function of rents and exogenous variables:

$$D\ (R, Exogenous\ Variables)\ =\ S \qquad (1)$$

The North-West quadrant has rents on vertical axis and prices on its horizontal axis. This quadrant depicts the ratio of rent to price, an exogenous capitalization rate for holding assets. This capitalization rate (i) is shaped by the long-term interest rate of the economy, any risk associated with the expected income flow, real estate taxation and the expected growth in rents. In this quadrant, the level of prices of the asset market (P) are determined by the level of rent (R) through the capitalization rate (i)

$$P = \frac{R}{i} \tag{2}$$

The South-West quadrant is where the supply of real estate market is decided. Has prices in the horizontal axis and construction stock (C). In this quadrant the construction costs curve ($\mathrm{f}(C)$) is drawn, pointing southwest. This suggests that if supply expands, construction costs expand as well. The degree of elasticity of the supply of real estate assets is related to supply constrictions, scarcity of land or any obstacle; the greater they are the more inelastic turns the supply curve. The level of prices of the market (P), already given, and the optimal level of construction costs determine the level of supply that ensures no excess of profits nor losses in the market.

$$P = \mathrm{f}(C) \tag{3}$$

Finally in the South-East quadrant, with construction stock (C) in the vertical axis and the stock for space in the horizontal axis. On this quadrant the constant flow of construction is yielded into real estate stock. Thereby, any variation in the real estate stock (ΔS) is a result of new stock reduced by any loss in stock adjusted by depreciation (d):

$$\Delta S = C - dS \tag{4}$$

The line drawn in this quadrant represent the levels of stock where the construction level replaces the loss in stock, keeping the level of construction stable over time.

$$\Delta S = 0 \quad \textit{and therefore} \quad S = \frac{C}{d} \tag{5}$$

Summarizing, the market for real estate space dictate rent levels. These rents through the market for real estate assets rule real estate prices. These real estate prices successively give rise to construction. The new stock resulting from it generates a new size of stock. Both, property, and asset markets reach the equilibrium when the level of stock is constant.

2.1.2 Comparative Statics

Comparative statics analysis study the signs of the changes in endogenous variables of the model as a result from changes in some underlying exogenous variables. Those exogenous variables cover phenomena outside the model such as the macro environment (unemployment, inflation, level of disposable income, long-term and short-term interest rates, gross domestic product trend, etc…), the demographic factors (population, number of households, etc…), the real estate taxation treatment or the availability to access sources of financing.

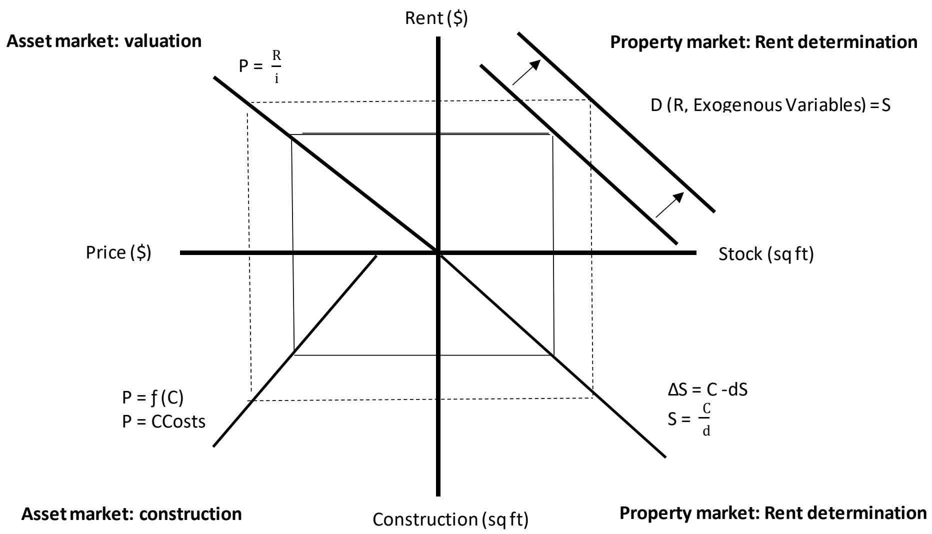

Figure 4. Property demand curve shifts upwards (adapted from DiPasquale and Wheaton (1992), p. 191)

In Figure 4, increases in the employment rate or the number of households expand demand for space, in North-East quadrant demand curve shifts upwards. In the short-term, with a given amount of real estate space, rents increase, the model predicts a positive relationship between demand and level of rents. In the North-West quadrant, higher level of rents increments real estate asset's value, which generates a higher level of construction (South-West quadrant). This increase generates a higher stock of space (South-East quadrant).

On the one hand, model predicts that rents, prices, construction level and stock of real estate will be higher after demand expands (economics growth), showing a positive relationship (sign) among these variables and the demand for real estate space and vice versa. On the other hand, the price-elasticity of each curve (demand for space, demand for real estate asset or construction cost) determines the magnitude of the changes. For instance, high price-elasticity of demand for space, for little changes in demand of space would increase the level of rents.

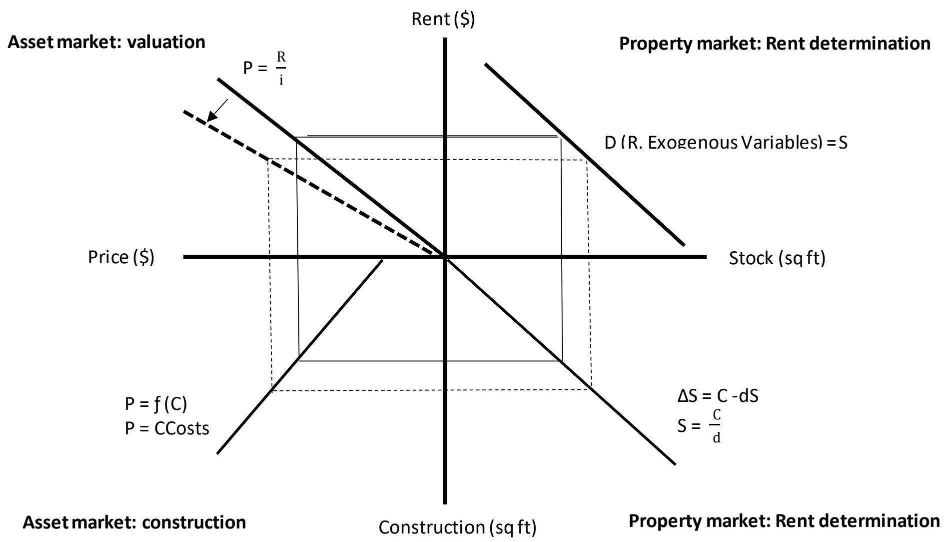

Figure 5. Asset demand rotates anti-clockwise (adapted from DiPasquale and Wheaton (1992), p. 193)

Demand for real estate assets can shift, due to changes in interest rates, changes in real estate tax treatment, or changes on how investors perceive risk of owning property. For example, more favourable tax regulations on real estate income, lower interest rates, or lower degree of risk rotates demand anti-clockwise, as it can be seen from Figure 5 (and vice versa). Given a level of rent, if expected income flows from real estate improve, rises the price for real estate asset. South-West quadrant shows how this rise in the price level expands construction. A higher level of stock of space is eventually lowering the level of rents.

A climb in demand for real estate assets, increases real estate prices, stimulates construction, and expands stock for space by pushing rents down (and vice versa). In addition, he negative relationship between interest rates and real estate prices is theorised in the model.

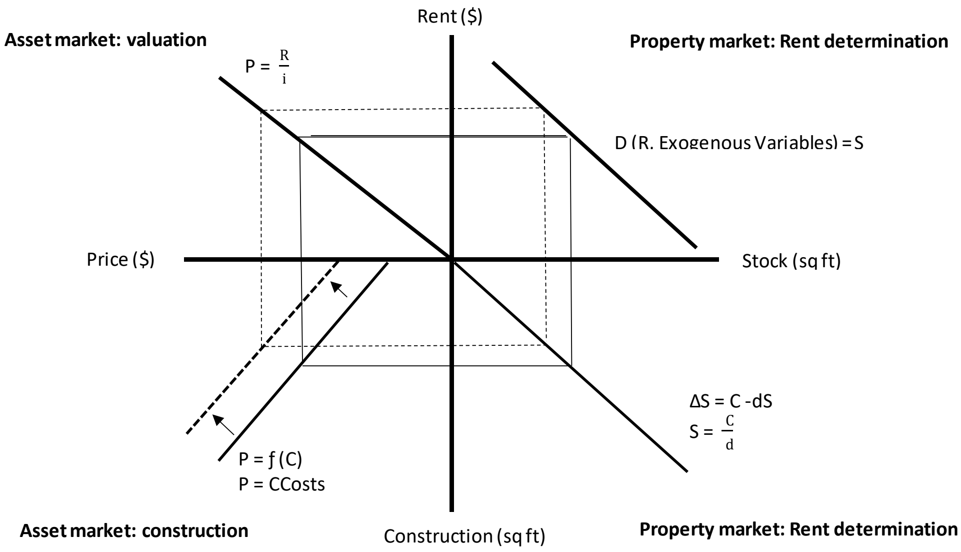

Figure 6. Asset costs shifts leftwards (adapted from DiPasquale and Wheaton (1992), p. 196)

Finally, the effect of changes in interest rates, the borrowing capacity of building contractors, or building regulations induces shifts in supply curve (construction curve) as shown in Figure 6. Upward movements of interest rates, increasing building regulations, or the difficulty to access credit due to credit rationing, reduce the profitability and therefore reduce the new building developments and contracts construction. For a given level of real estate prices, a negative shift in supply, lowers the level of construction, reducing the stock of space and eventually increasing rents. The latter give rise to real estate assets' price to rise.

2.2 DiPasquale and Wheaton (1994): the stock-flow model revisited

Stock-flow macro-model begin with a two-equation model with supply (S) and demand functions (D). The demand for housing (6) depends on a set of exogenous variables (X_1), a set of fundamentals affecting demand such as permanent income, demographic factors, etc…, the real house prices (P), the real user cost of owning a dwelling (U) and the cost the owner would face if did not own the place of residence (R). The supply function (7) is a differential equation which depends on the level of construction and the amount of stock depreciated at rate δ (δS). At the same time construction function depends on some exogenous variables (X_2), such as production costs, interest rates, the real house prices (P).

$$D\ (X_1, P, U, R)\ = S \tag{6}$$

$$\Delta S = C(X_2, P) - \delta S \tag{7}$$

The classic assumption of stock-flow model is summarized in equations (6) and (7), where equilibrium prices adjust rapidly allowing demand and supply to adjust. Since market adjust quickly and clears, demand equals the current housing stock at any time. Therefore, the equilibrium real prices are, at any time, shaped by fundamentals affecting demand and supply, as in equation (8):

$$P = \mathrm{f}\ (D, S, U, R) \tag{8}$$

DiPasquale and Wheaton (1994) consider some modern approaches on how the endogenous variable U (real cost of owning a dwelling) is to be found on some equation resembling (9)[4]:

$$U = \left(i + t_p\right)\left(1 + t_y\right) - E\ (\frac{\Delta P}{P}) \tag{9}$$

where $\left(\mathrm{i} + \mathrm{t_p}\right)\left(1 + \mathrm{t_y}\right)$ represents the after-tax cost of debt and property taxes. Being i, the nominal interest rate; $\mathrm{t_p}$ is the marginal tax rate on real estate; $\mathrm{t_y}$ is the marginal tax rate on capital income and $\mathrm{E}\ (\frac{\Delta \mathrm{P}}{\mathrm{P}})$ represents the real housing appreciation or the average real housing price growth. Equation (9) is inserted in equation (6).

The new approach from DiPasquale and Wheaton (1994) debates some of the assumptions of the stock-flow model by bringing in a more realistic perspective. It stresses on four points. First, on why it is not realistic to sustain real estate market clearance. Secondly, adaptative expectations are included in the model, setting a new perspective on how consumers shape their expectations on future house prices. Thirdly, debates the relationship between construction of new housing stocks and

[4] Later, we go deep into this issue with Poterba (1984) and Himmelberg et al. (2005)

house price levels, and construction and availability of land. Finally, on how demographic developments influence real house prices and housing stock growth.

The assumption of market clearing is relaxed; thus, inefficiency of housing market describes better why market prices are much sluggish than expected and why real estate market responds slower than theory predicts (product heterogeneity, time-consuming search, etc…). Moreover stock-flow model predicts that supply will steadily adjust, believing that demand side will adjust emptying the market completely. That is not realistic, it is to be expected that demand and price gradually adjust lead to gradual market clearing. This assumption is included in the stock-flow model with some price adjustment mechanism:

$$P_t = \tau P_t^* + (1 - \tau) P_{t-1} \tag{10}$$

where P_t^*, represents the theoretical equilibrium price P_t; represents the level of prices, P_{t-1} is the one-period-lagged price; τ is a convergency rate, telling about how fast prices converge towards the theoretical equilibrium.

At this point, adaptative expectations are incorporated into the model. Although rational expectations would be preferred, they believe that agents form their expectations by looking backwards. This way, prices keep an intelligible form and show serial correlation with their lagged values (direct intertemporal price correlation).

DiPasquale and Wheaton (1994) give a new role to land, as a driver of supply of housing. Previous models consider that supply reacts permanently to rise in prices by increasing stock of housing. Rising costs in construction are in the short run mostly caused by congestions in the production process and in the long run by the availability of production factors (labour and resources). The price of land correlates with the level of activity. But now, the price of land depends on the existing stock of housing instead. When prices increase give birth to “excess returns”, boosting construction. As the level of housing expands, the price of land increases and exhausting the excess returns, by increasing construction costs. Hence supply increases only transitorily until the stock of housing matches with the new long-term equilibrium stock.

It is possible to draw a new supply of housing stock that connects these developments of the stock of housing (considering lagged stocks) and the long-term equilibrium housing stock, as in (11):

$$\Delta S = C - \delta S = \alpha[\, S^{*}(X_2, P) - S\,] - \delta S \tag{11}$$

where $S^{*}(X_2, P)$ represents the long-term supply (stock of housing), S is the current stock of housing, α, is the rate to at which the housing stock converges. Hence, housing prices determine the construction of new housing stocks if S^{*} is higher than the current stock of housing.

Based on spatial theory, DiPasquale and Wheaton (1994) say that the stock of housing is expected to grow at a similar pace than population grows. Demographic developments, through demand for housing, drive the supply of stock and therefore has a part in the long-term price determination. Is in the words of DiPasquale and Wheaton "in case of price elastic construction in the short-run and the rising long-run supply curve for the stock of housing, it is impossible for prices to undergo any sustained decline" (p. 2). This way, under normal circumstances demographic trends can reduce the growth in real house prices, but those are not likely to fall.

2.3 Poterba (1984): an asset-market model of the housing market

The original aim of this model was of analysing the impact of tax policy and inflation on the relative price of houses and on the size of the housing capital stock (Poterba, 1984, p. 730). More to the point, about how changes in the real user cost have effects in the short run and the long run dynamics of housing market.

According to Poterba, sudden changes of house prices are conditional on future expectations on construction activity. Moreover, he assumes that households correctly interpret the effects of such variations (equalising the price of the property with the present discounted value of its potential flow of money (Poterba, 1984, p. 730). This way, future flows will depend negatively on the growth of housing stock (as the user costs of ownership decrease as stock gets bigger).

The quantity demanded of housing services,[5] HS^d, is dependent on the real rental price, thus (12) represents the demand in the housing market:

$$HS^d = f(R), f_R < 0 \tag{12}$$

The supply of the market, HS^S , is determined by the construction firms, H, which supply a homogeneous service to the market; Hence the supply of the market is:

$$HS^S = h(H) \tag{13}$$

Therefore, the demand (12) and the supply (13) of housing services equal, as in (14):

$$HS^d = HS^S \tag{14}$$

In the short run, the stock of housing is fixed, and the real rental price acts as a market-clearing condition. R depends on the existing stock of housing in the market, with negative first derivative (decreasing with stock), as:

$$R\ (H) = R\ (h(H)),\ R' < 0 \tag{15}$$

According to economic theory, households are willing to consume housing services until their marginal value equates marginal costs.

Hence, the (one-period) user costs of housing services (ω) can be calculated as a sum as in (16): if the housing stock depreciates at a rate δ (constant) and involves some preservation payments κ[6]; if households sustain property tax liabilities at a percentage μ and income tax rates on capital gains at a marginal tax rate, θ[7]; if households can

[5] We comply with Poterba's original nomenclature.
[6] Represents a fraction of current value of the stock of housing owned by households.
[7] Model assumes that property tax can be deduced from dutiable income. In Finland capital gains from owner-occupied dwellings are tax-exempt.

lend or borrow money at interest, i[8], understood as well as the cost of opportunity of owning real estate instead of other assets; and finally, π_H, the rate of capital gain. Thus,

$$\omega = [\,\delta + \kappa + (1 - \theta)\,(i + \mu\,) - \pi\_H\,] \tag{16}$$

If Q represents the real house price, and ω is the real user cost, then ωQ represents the (one-period) real cost of ownership, in terms of the cost of house value. Therefore, the real cost of ownership has to equal (15):

$$R\,(H) = \omega Q \tag{17}$$

Equation (17) is the "arbitrage condition" of the housing market (between owning and renting). Arbitrage condition helps to evaluate whether house prices are above their level (or below). For instance, if real user costs rise by keeping rent unchanged, "arbitrage condition" forces prices to fall (and vice versa).

2.4 Himmelberg, Mayer and Sinai (2005): the user cost of housing

According to Himmelberg et al. (2005), the financial return of an owner-occupied estate compares the "imputed rent", a measure of the cost to lease such estate, with the "opportunity cost of capital", the loss or gain of revenue from capitalising wealth in an alternative investment.

Thus, the user cost of housing (or homeownership) is a reasonable way of assessing housing prices. This user cost (annually calculated) helps to compare whether the cost of owning is or not disproportionated with the cost of renting.

Himmelberg et al. (2005), review the components of homeownership: the first component is the one-year (opportunity) cost of the sacrificed interest that proprietor could have earned by investing in some alternative investment. This cost is calculated as the price of housing at the period t, P_t, multiplied by the risk-free interest rate, r_t^{rf}.

[8] Represents the mortgage interest rate.

An Additional component is the (annual) cost of property taxes[9], computed as the house price, P_t times the property tax rate, ω_t. Another one is the tax deductibility of mortgage interest and property taxes, calculated as the applicable tax rate on income multiplied by the mortgage and real property tax expenditures, $P_t\ \tau_t\ (\ r_t^m +\ \omega_t)$. The next element reflects the maintenance costs as a portion δ_t of dwelling value. The second last term, g_{t+1}, is the anticipated capital gain (or loss). Finally, the last element is a risk premium component, to compensate homeowners as opposed to renting, $P_t\ \gamma_t$.

Thus, adding all the components we get the annual cost of homeownership at period t, CO, as:

$$CO = P_t\, r_t^{rf} + P_t\omega_t - P_t\tau_t\,(r_t^m + \omega_t) + P_t\delta_t - P_t g_{t+1} + P_t\gamma_t \qquad (18)$$

Therefore, the equilibrium in the housing market is found when the annual cost of homeownership equals the annual cost of renting R_t, equivalent as (17) in Poterba (1984).

$$R_t = P_t\, u_t\ {}^{10} \qquad (19)$$

In (19), the fraction u_t is the user cost of homeownership (or user cost of ownership). Moreover (19) states that if cost of ownership decreases (increases), by keeping rents, then price level has to increase (decrease). This correcting process is in the same vein than Poterba (1984), this is according to Himmelberg et al. (2005), the "no arbitrage" condition. Rearranging terms we get:

$$u_t = r_t^{rf} + \omega_t - \tau_t\,(r_t^m + \omega_t) + \delta_t - g_{t+1} + \gamma_t \qquad (20)$$

[9] In Finland capital gains from owner-occupied dwellings are tax-exempt.

[10] Reorganizing (19), the inverse of the user cost turns into the price-rent ratio, P_t/R_t. This ratio benchmarks renting against owning property and it is widely used to assess whether markets are fairly prized: $P_t/R_t = 1/u_t$.

where u_t, is expressed in terms of cost "per currency" of house value. Lower real interest rate reduces user cost since the cost of opportunity of investing in alternative assets and house ownership becomes more desirable (since mortgage interest is tax deductible). Increases in income tax rates, has a similar effect on user cost, makes buying more attractive than renting. Increases in the risk-free interest rate or the maintenance costs rise the user cost, makes renting more attractive than buying property. Finally anticipated gains of owning a property will decrease user cost, making more attractive buying than renting.

2.5 Meen (1990): the life-cycle approach

The user cost in (20) is in essence, similar to the user cost in Meen (1990). Meen uses the life-cycle model as point of departure. This is a multi-period consumer utilization problem[11] with two commodities, housing (H)[12] and an amalgam of other consumption commodities (C).

The life-time utility in continuous time is:

$$\int_0^\infty e^{-rt} \ \mu\big(H(t), C(t)\big)dt \tag{21}$$

where r represents the real discount rate and $\mu\big(H(t), C(t)\big)$ embodies the utility function. Equation (22) represents the household's budget constraint at period t:

$$g(t)X(t) + S(t) + C(t) = (1-\theta)Y(t) + (1-\theta)i\,A(t) \tag{22}$$

The left-hand side of the equation refers to all combinations of goods households that households can afford given the prices[13] and the savings: $g(t)$ represents the real house price at t; $X(t)$ is the purchase of housing and therefore $g(t)X(t)$ the value of housing

[11] Financial market has no restrictions and lending and borrowing proceeds at the interest rate, i, exogenously given.
[12] Meen assumes that the flow of housing services is proportional to the housing stock and an argument in the utility function.
[13] Consumption commodities (C) are treated as the numeraire: its price is normalized.

consumed at period t; $S(t)$ are the real savings at period t; and $C(t)$ is the amount of non-housing consumed at period t. In the right-hand side of the equation, $(1-\theta)Y(t)$ is the real disposable income at period t discounted by the marginal income tax rate, θ; the second component $(1-\theta)i\ A(t)$ stands for the other assets that household have at period t discounted by the post-tax return $(1-\theta)i$.

The evolution of the stocks of other assets ($\dot{A}(t)$) that households hold overtime is represented by:

$$\dot{A}(t) = S(t) - \pi\ A(t) \qquad \textit{(23)}$$

In (23) the asset stock held by family units is relative to the level of household's savings adjusted by inflation; thus, π indicates inflation rate and hence characterizes the depreciation rate of other assets household have at period t.

The net change in housing stock $\dot{H}(t)$ is characterized:

$$\dot{H}(t) = X(t) - \delta\ H(t) \qquad \textit{(24)}$$

where δ is the depreciation rate of housing stock. In equation (24) housing stock evolves over time according to the level of new acquisitions of housing adjusted by the depreciation of the existing housing stock.

The maximisation of the utility function, $\mu\big(H(t), C(t)\big)$ with respect to (22) derives the marginal rate of substitution[14] between housing and the amalgam of other consumption commodities ($^{\mu_h}/_{\mu_c}$). This is the real user cost of capital[15]:

[14] From the first order conditions.

[15] Meen (1990) presents an alternative form to (25) including a term : the shadow price of credit rationing to the marginal utility of the amalgam of other commodities: $^{\mu_h}/_{\mu_c} = g(t)[(1-\theta)i - \pi + {}^{\lambda(t)}/_{\mu_c} + \delta - \dot{g}/g\ (t)]$. Alternatives forms include some risk premium as well, γ (Oikarinen, 2009).

$$^{\mu_h}/_{\mu_c} = g(t)[(1-\ \theta)i - \ \pi + \delta - \ \dot{g}/g\ (t)] \tag{25}$$

In equation (25), $\dot{g}/g\ (t)$, embodies the expected real capital gain of owning housing.

Assuming that within housing market coexist two interdependent markets, the market for real estate and the market for real estate assets. In the market for real estate, demand and supply for housing determine the real imputed rental price of housing services ($R(t)$) that clears the market. The demand of housing depends on the disposable income, on some demographic variables and on the $R(t)$. Under Meen (1990) assumptions, the supply of housing is rigid in the short run.

In the second market, given $R(t)$, the real house price for assets ($g(t)$) has to adjust to keep market in equilibrium. Therefore, market efficiency requires that arbitrage relationship in equation (17) holds. Rearranging (25) and including expectations:

$$g(t) = \frac{R(t)}{(1-\ \theta)i-\ \pi^e+\delta-\ \dot{g}^e/g(t)} \tag{26}$$

Equation (26) characterizes the real price equation.

According to Meen (1990), $R(t)$ it is not directly observable and has to be proxied, therefore for empirical purposes it is replaced by its determinants. He assumes that real rental price is a function on income, demographic variables, and the current stock of housing. Other authors consider other determinants as well, such as real household wealth or the real interest rate (see e.g., Oikarinen 2007).

2.6 Literature review

Modern empirical research concerning house price fundamentals lay emphasis on two main questions: the long-term house price equilibrium and the short-term dynamics by identifying deviations from its long-term equilibrium levels.

A widespread range of fundamentals of house prices are found in empirical literature such as disposable income, construction costs, demographic and labour factors or user cost of ownership (interest rates, rental yield of a property, mortgage interest deductibility), the access to credit or institutional factors, among others.

Some of the literature studies housing markets developments at different levels, mainly at a metropolitan or regional level (see e.g., Case & Shiller 1990, Poterba et al. 1991, Abraham & Hendershott 1996, Chen & Patel 1998, Hort 1998, Kuismanen et al. 1999, Malpezzi 1999, Capozza, Hendershott, Mack & Mayer 2002, Meese & Wallace 2003, Oikarinen 2005, Oikarinen 2009, Oikarinen 2014, Bourassa, Hoesli & Engblom 2018, Bourassa et al. 2019) but as well at national level (e.g.,, Poterba 1984, Topel & Rosen 1988, Mankiw & Weil 1989, Meen 1990, Drake 1993, DiPasquale & Wheaton 1994, Barot & Takala 1998, McCarthy & Peach 2002, Jacobsen & Naug 2004, Taipalus 2006, Hott 2009, Gimeno & Martínez-Carrascal 2010, Steiner 2010, Wang & Tumbarello 2010, André & García 2012 or Oikarinen 2012).

Extensive literature covers the short-term dynamics as well, especially on those forces (driven by momentum). Lagged variables enter the equation as explanatory variables, embracing the idea of housing market as an inefficiency market. Hence, exists broad consensus on some degree of price "sluggishness", prices seem to move slower than it would be expected towards their long-term equilibrium levels. (See e.g., Drake 1993, DiPasquale & Wheaton 1994, Barot & Takala 1998, Malpezzi 1999, Oikarinen 2005, Oikarinen 2009, Steiner 2010, André & García 2012 or Oikarinen 2014).

In the present years, the availability and quality of data (micro-level data) and the improvement in modelling techniques infused qualitative improvements in research. Heterogeneous and imperfect data, time-varying structural attributes, or regional specific features conditioned differences in empirical results. Most of studies use quarterly data (see e.g., Kearl 1979, Poterba 1984, Topel & Rosen 1988, Case & Shiller 1989, Case & Shiller 1990, Meen 1990, Takala & Pere, Drake 1993, Zhou 1997, Barot & Takala 1998, Chen & Patel 1998, Meen 2002, McCarthy & Peach 2002, Jacobsen & Naug 2004, Oikarinen 2005, Oikarinen 2006, Oikarinen 2009, Gimeno & Martínez-Carrascal 2010, Wang & Tumbarello 2010, André & García 2012, Oikarinen 2012, Oikarinen 2014, Oikarinen et al. 2018 or Bourassa et al. 2019) but some use yearly data[16] (see e.g., Mankiw & Weil 1989, Poterba et al. 1991, DiPasquale & Wheaton 1994, Abraham & Hendershott 1996, Hort 1998, Kuismanen et al. 1999, Malpezzi 1999[17], Capozza et al. 2002, Capozza or Hendershott & Mack 2004). Meese and Wallace (2003) preferred monthly data.

Most of current literature embraces dynamic perspective studying both long-term and short-term causalities. Models with mean reversion, such as VECM[18] or ECM, and cointegration tests have become recursive approaches in house price research (see e.g., Takala & Pere 1991, Chen & Patel 1998, Hort 1998, Malpezzi 1999, Meen 2002, McCarthy & Peach 2002, Mees and Wallace 2003, Oikarinen 2005, Oikarinen 2009, Gimeno & Martínez-Carrascal 2010, Wang & Tumbarello 2010, André & García 2012, Oikarinen 2012, Oikarinen 2014, Oikarinen et al. 2018 or Bourassa et al. 2019).

Cointegration tests are used to identify long-term interactions between two (or more) variables. Cointegration has implications in long-term dynamics: two (or more) non-stationary variables are expected to have some long-term relationship with each other. One of the first econometric research to incorporate Engle-Granger cointegration analysis in U.K. was Giussani and Hadjimatheou (1990). Engle-Granger test (Engle & Granger, 1987) for cointegration suits to study one equilibrium relationship with a testing procedure based on ECM. Zhou (1997) forecasts sales and price for homes in U.S. using Engle-Granger methodology. Hort (1998) estimates an ECM using an Engle-Granger two-steps procedure. In the first step estimating the long-term relationship and in the second step she explores the short-term dynamics through the residuals from cointegration regression. Malpezzi (1999) with a simple ECM looks for long-term relationships among house prices and some other variables, such as income. André & García 2012 estimate jointly housing prices, residential investment, and housing stock in Finland over the period 1995-2010. The short-term relationship is

[16] Most carry panel data analyses (e.g., Capozza et al. 2002, Capozza et al. 2004, Host 1998 or Oikarinen et al. 2018).
[17] Data is quarterly but converted to annual changes.
[18] Vector Autoregressive Models (VAR), Factor-Augmented VAR Models or Bayesian Models are employed as well as empirical models, but they are left aside, since the scope of this master thesis is to focus on VECM and ECM.

estimated in Granger-Engle ECM. The long-term connection is estimated with a system of equations weighted by Two-Stages Least Square (2SLS).

Later Drake (1993) incorporates to avoid small sample bias the Johansen cointegration test (Johansen 1988, Johansen & Juselius 1990). Johansen cointegration test or maximum likelihood estimation (ML) is suitable for multivariate cointegration. Johansen test method is grounded in VECM. Drake estimates the unknown parameters of his regression model using maximum likelihood estimator. Takala and Pere (1991) test the cointegration house and stock prices in Finland. They use the Johansen procedure to estimate a VECM. Chen and Patel (1998) examine the dynamic relationship between house prices and their determinants in Taipei[19]. They use the two-stages procedure, estimating the cointegration matrix by Johansen procedure. Barot and Takala (1998) cointegrate house prices and inflation for Sweden and Finland. The method selected to estimate parameters is the full information maximum likelihood. Oikarinen (2012) estimates a VECM, to study the co-movement of house prices and transactions in Finland, by Johansen maximum likelihood method. Gimeno and Martínez-Carrascal (2010) study the interdependencies between house prices and mortgages in Spain, they estimate a VECM using Johansen maximum likelihood estimation in the same vein than Oikarinen (2009).

Literature covering regression analysis to detect long-term relationships among variables can be classified broadly in different groups: some authors assess price levels against supply and demand variables considering a consumption good (see e.g., DiPasquale & Wheaton 1994, Abraham & Hendershott 1996[20], Malpezzi 1999, McCarthy & Peach 2002[21], Riddel 2004, Oikarinen 2005, Steiner 2010, Oikarinen et al. 2018 or Bourassa et al. 2019); an additional set of authors study house prices as the present value of future flows, with models based on asset price concepts (see e.g., Kearl 1979, Poterba 1984, Topel & Rosen 1988, Mankiw & Weil 1989, Poterba et al. 1991, Takala & Pere 1991, Barot & Takala 1998, Chen & Patel 1998, Hort 1998,

[19] Republic of China, Taiwan.
[20] They estimated a two-equation model, one for the long-term equilibrium and another for adjustments to the equilibrium.
[21] They implement the DiPasquale and Wheaton stock-flow model in an ECM framework.

Capozza et al. 2002, Jacobsen & Naug 2004 or Hott 2009); other researchers focus on the life-cycle model, model derives the marginal rate of substitution (MRS) between accommodation and an aggregate of consumption commodities (see e.g., Meen 1990, Meen 2002, Oikarinen 2009, Oikarinen 2012 or Oikarinen 2014); some additional group conduct cointegration analysis (see e.g., Drake 1993, Zhou 1997, Oikarinen 2006, Gimeno & Martínez-Carrascal 2010); finally the approach relying on unit root tests (see e.g., Taipalus, 2006).

Literature on the house market inefficiency is large. Case and Shiller (1988) tackle the problem of the efficiency of the market of private housing. Their study covers a time span starting 1970 to 1986, by employing sales price indices for 4 different cities from United States. They notice that housing prices do not follow a random walk and therefore, are forecastable. Although forecasting only works for those prices showing positive serial correlation. Another conclusion of their research is that housing market is way less efficient than financial market, where excess returns have persisted long time. They identify some inefficiencies that may lead to such conclusions such as high transaction costs, carrying costs, tax considerations or no profit opportunities when prices are declining. Similar result about the inefficiency of the housing market is found in Case and Shiller (1990) or in Mankiw and Weil (1989) who believe the housing market it is not an efficient market, since housing prices do not seem to reflect arbitrage conditions. Other authors observe similar results regarding house market inefficiency (see e.g., Hamilton & Schwab 1985, Linneman 1986, Rayburn, Devaney & Evans 1987).

Mankiw and Neil (1989) highlight the paper of demographics in shaping both, the demand for housing and housing prices, although its sign with respect to house prices it is disputable, they expected to be positive. As a result from “baby boom”, more households were created over the period 1970-1980, shifting demand for housing (slowly). Based on birth projections, they predict that once the demographic effect vanishes, the housing prices will fall (as of 1990s and onwards). In the same vein, Kuismanen et al. (1999) base their empirical analyses on the Mankiw-Weil approach. They study the role played by of demographics in the housing market of the regional rea of Helsinki. They find a positive relationship between house prices and demographics: their results show that about 1% increase in population increases about

0.2% the house prices. DiPasquale and Wheaton (1994) respond to Mankiw and Weil, that even though demographic determinants may have an influence on housing prices, the effect of demography can only slow the escalation of prices, but not lead prices to fall (if income does not decline). Case and Schiller (1990) make use of some adult population variable, when forecasting house prices and excess returns. They uncover a positive a relationship between prices and population growth. Poterba et al. (1991) try a similar approach for 1980s at US city-level, but without identifying any significant relationship between the evolution of prices and the demographic demand variable. Hort (1998) utilizes a sample of adult population aged 25-44 years to capture the effect of age over house demand. Oikarinen (2005) uses a more restricted sample of population aged 20-29 years, since according to him, people falling between 20 and 29 show more mobility and their housing consumption increases faster.

Another implication from Mankiw and Weil (1989) is to incorporate some “age composition of population” proxy, segmented by cohorts of population, when approaching demand for housing. Of these cohorts of population, those from 20 to 30 years old show strong positive demand of housing while in cohorts over 40 years old the quantity demanded for housing decreases. Many authors incorporated this feature on their research with little variations. For example, Abraham and Hendershott (1996) consider population aged from 25 to 64 when adjusting the income per working adult. Kuismanen et al. (1999) construct demographic time series clustering CRH population in 5-years age groups. Jacobsen and Naug (2004) incorporate shares of population from 20-24 and 25-39 through wage income. In André and García (2012), who study the price and investment dynamics in Finland, includes a demographic variable (population from 25 to 44 years) and finds a positive relationship between size of population and house prices.

Some authors have focused their attention on the effect and existence of speculative bubbles in the housing market. Abraham and Hendershott (1996) study the existence of housing market bubbles in U.S. metropolitan areas with yearly data over the period 1972-1992. They shape a model to account for a market that is constantly adjusting towards an equilibrium price level (including time lags in the adjustment of prices). The existence of two separate groups of variables is highlighted, one comprising the main determinants of house price changes (real income growth (0.57), the real construction costs (0.46) and the real after-tax interest rate (-0.59) and another with those describing the adjustment dynamics from equilibrium prices (with the lagged real appreciation rate ("bubble builder" of 0.4) and the difference between the real house price level ("bubble burster" of -0.07). According to Abraham and Hendershott, three-fifth of the variation is explained by all variations, although over two-fifths by each of the variable groups. Most of U.S. territory seems to be in near equilibrium, although they detect some overpricing by 1992 between 15 and 30 percent in Northeast and West Coast areas.

With monthly data Taipalus (2006) conducts unit root tests for the log rent-price ratio for Finland, Germany, Spain, U.K., and some U.S. cities. This way tracking the phases when housing prices diverge from their essentials. For instance, for Helsinki from autumn 1987 until the first part of 1989, she uncovers a possible bubble. Although, rent controls in force at the end of 1980s make rent increases difficult "pari passu" with prices.

Oikarinen (2005), over the period 1975Q1-2005Q2, explores the long-term relationship between house prices and their fundamentals for the Helsinki urban area. He aims at finding whether house prices are experiencing or not a bubble. According to Oikarinen, house prices are not overpriced at the end of the period, although he detects some overvaluation in some downtown areas. André and García (2012) believe that some determinants in Finland can cause volatility in the housing market[22], even in the absence of a bubble (p. 8).

Hott (2009) suggests in his own words that "house prices fluctuate more than fundamentally justified" (p. 1). He assumes agents' rationality with perfect foresight as starting point. He calibrates an asset pricing model for 6 countries with adaptative expectations (based on constant user costs) and tests three of the feasible explanations for price volatility: speculative bubbles, momentum trading and herding behaviour (p. 16). Results are somehow inconclusive: momentum trading seems to fit for Ireland; The speculative bubble is a good determinant for U.S., The Netherlands and Japan and

[22] Such as loan rates or household income.

the herding behaviour seems to fit better for U.K. and Switzerland (p. 24). There is a flaw in the model: financial system is not included in the model, and therefore are not the effects of financing (access to mortgages) on demand.

On the other hand, Bourassa et al (2019) perform a different approach, firstly quantifying house price bubbles with an asset pricing approach to subsequently compare the results with other methods such as price-rent ratio analysis, multivariate regression, or imputed-actual rent ratio. Bourassa et al. use quarterly data from for a set of 6 metropolitan areas that have suffered substantial increases in prices, covering a time span over three decades (Helsinki from 1975Q1 to 2012Q4). They aim to evaluate models' effectiveness (ex-post but also in identifying bubbles), surprisingly the price-rent ratio approach performed better than other sophisticated methods with an effectivity of 88.6% recognising ex-post bubbles and 84.1% identifying imminent bubbles.

Income[23] variable plays a significant role shaping house prices. Most of the existing literature agrees on its prevailing effect. Income elasticities are estimated in literature under different ways: per capita income (see e.g., Drake 1993, Hort 1998, Malpezzi 1999), per household income (see e.g., Meen 1990, Takala & Pere 1991, Chen & Patel 1998, Oikarinen 2005 or Gimeno & Martínez-Carrascal 2010), aggregate income (see e.g., Poterba 1984, Oikarinen 2009 or Oikarinen 2014) or the income per adult (over 20 years old age population) as in Kuismanen et al. (1999). McCarthy and Peach (2002) use data related to consumption as a proxy or revenue as Meese and Wallace (2003).

Income's sign with respect to house prices is expected to be positive with respect to house prices. A group of authors report income elasticities under or close to 1. Mankiw and Weil (1989) obtain estimated income elasticities ranging from 0.234 and 0.26 (table 2); In the same vein, Case and Schiller (1990) estimating house prices and excess returns, find income elasticity to house price with coefficient 0.31 (p. 268). In Abraham and Hendershott (1996) income is the main driver of house prices in U.S.,

[23] Either total or in per capita terms:

with an income elasticity to price of 0.774 (table 3). Hort (1998) finds income as main determinant of long-term equilibrium house price for Sweden (with data from 20 cities for the period 1968-1994) with income elasticity close to unity, with coefficient 0,965 (eq. 3, p 105). Kuismanen et al. (1999) find, as expected, that house prices increase with income. They employ a lagged income variable with income elasticity of 0.808 (p. 28). Oikarinen (2005, p. 19) obtains comparable results. Income plays a key role determining long-term equilibrium house prices in the metropolitan region of Helsinki, with income elasticity ranging from 0.83 to 1.34. In the same vein in André and García (2012), Finnish house prices respond positively to income per capita, although coefficient is difficult comparable (2.65). Although later Oikarinen (2005) uses some “corrector variable” to capture the effect of financial liberalization: current income seems to be more determinant before the financial liberalization, and therefore coefficients are higher). After the adjustment, some income elasticities range from 0.3 to 0.47. Oikarinen (2009) estimates an income elasticity of 0.356 (p. 133). Wang and Tumbarello (2010) notice a positive relationship with respect to earnings/income with a coefficient of 1.20 in Australia (p. 11). Oikarinen at al. (2018) study the house price dynamics for 70 urban areas in the U.S. for the period 1980Q1 until 2015Q2. They find substantial city-level differences in long-term income elasticity to price, although in average is 0.81 (p. 23). Bourassa et al (2019) detect for Helsinki, Geneva, Zurich, Chicago and San Francisco positive and strong relationship between the aggregate income and house prices with income elasticities in a range from 0.44 to 1.82; Helsinki, 0.91 and San Francisco 1.1 (p. 551). Surprisingly, income is not statistically significant when determining house prices in Zurich. Other authors find comparable results (see e.g., Capozza et al. 2004 (0.45, p. 17), Jacobsen & Naug 2004 (1.22, p. 35)).

In the same vein than Case and Shiller (1990), Poterba et al. (1991) using U.S. city-level data over the period 1980-1989 intend to explain house price variations. They set a cross-section model to capture how price levels react to changes in some fundamentals. They find that changes in real income and construction costs are the main drives in price changes at U.S. city-level, but by contrast, their estimate an income elasticity over 2, 2.054 (p. 168).

Another group of authors, in a similar vein than Poterba et al. (1991) which estimate high income elasticities. Meen (1990) observes for U.K., that under mortgage rationing, the elasticity of prices to income is quite high, 3.0 (p. 17), having considerable influence on house prices. Drake (1993) for U.K. gets as the driving force behind variations of house prices, with income elasticity of 3.8, comparable to what Meen (1990) obtained (p. 1227). McCarthy and Peach (2002) for U.S., they use a consumption growth as a proxy or income, their results are consistent with theory and the sign as expected. The income elasticity is about 3.4 (p. 146). Other authors find similar results (see e.g., Meen 2002 (U.K. 1.18-2.54; U.S. 2.71, p. 10), Riddel 2004 (2.77, p. 127)).

Another determinant of the house prices is the construction costs. These are usually included in the empirical analysis. According to DiPasquale and Wheaton (1992), when short-term prices and construction costs differ, construction activity emerges (if prices are above construction costs) and therefore increasing construction costs. The expected sign between house prices and construction costs is positive. Poterba et al. (1991) observe that along with real income, construction costs are the second main driver in price changes. Ordinary Least Squares (OLS) estimation yields values of coefficients about 0.97: one percentage point increase in construction costs causing real prices to change around 0.97% (p. 163).

Other authors who examined the explanatory power of real construction costs are Case and Shiller (1989) and Case and Shiller (1990), who find construction costs having the expected value, positive with respect to prices (positive with coefficients ranging from 0.122 to 0.225 depending on the city) (Case and Shiller 1990, p. 268). Chen and Patel (1998) claim that construction costs have, as predicted, a positive effect on prices. They account about 10% of the variability of future prices in Taipei housing market. Abraham and Hendershott (1996) estimate an elasticity of construction of 0.35 for the whole U.S., although in coastal cities this elasticity is 0.163 and in interior areas 0.575 (table 3). In Hort (1998) the elasticity of prices to construction in the preferred model (eq. 3) is positive and 0.50, although results among models is not fully robust since they range from 0.269 to 0.583 (p. 105). Malpezzi (1999) finds that rigorous regulations relate positively with prices in some U.S. cities. Hence, increasing construction costs. Meen (2002) says that in most British studies, rarely construction cost enters the equations as regressor whereas it is more common to find construction costs in U.S. studies (p. 11). Meese and Wallace (2003) for Paris, observe that the sign of construction costs is according to the theory, although the magnitude may make difficult comparisons (coefficient of 6.57) (p. 1034). Capozza et al. (2004) conduct panel data analysis from 62 metropolitan areas in U.S.. Using median values, their model (eq. 1) reveals that a 0.01 unit change in the construction cost leads to a 1.1 percent increase in house prices, in the line of the theoretical prediction of 1.03 at 5% level significance (p. 16).

In Oikarinen (2005) construction costs range from 1.10 and 2.27, depending on the model, all of them positive as expected (p. 19). Although, later in Oikarinen (2009, p. 130) says that statistics do not support the inclusion of any supply-side variable either in the short-term or the long-term model. Riddle (2004) gets no significance of construction costs as a house price determinant, neither in the long run or the short run. She finds that in the short-term construction costs have no effect on prices, (0.001, pp. 129-130), although she recognises that the model may have some misspecifications, with unexpected results such as the negative relationship between rents and house prices. In Bourassa et al. (2019) construction costs have positive sign, as expected, for all the 6 cities. Although, they are only statistically significant for 4 of them (For Helsinki, the construction elasticity to price, 0.283, is in the same direction than Oikarinen (2009) and not statistically significant. Neither for Zurich) (p. 551).

The real user cost of homeownership it is also included as a determinant of house prices in many empirical research papers, according to theory its sign is expected to be negative with respect to house prices. Its correct calculation entails difficulties and sometimes its components are proxied independently (see e.g., Oikarinen 2009 and 2012) where the loan-to-GDP ratio proxies for the shadow price of mortgage rationing divided by the marginal utility of consumption ($^{\lambda_t}/_{\mu_c}$). Different ways are found in literature to approach its value (e.g., compare Poterba 1984, Meen 1990, DiPasquale & Wheaton 1992 or Himmelberg et al. 2005). Other authors (see e.g., Hort 1998, Oikarinen 2005) do not include the depreciation and maintenance rate (δ) since they consider it constant along the studied period, or do not consider the risk premium (γ) in it (e.g., Meen 1990 or Meen 2002). Poterba (1984) believes that a combination of the tax system and inflation were key factors to understand why during the 1970s tenure became desirable in the U.S.: the deductions applied for home loan interests and the exemptions from real estate gains along with high inflation rates, reduced the real user cost of ownership households were facing. As a result, house prices increased sharply. In Case and Schiller (1989) and Case and Schiller (1990), changes in marginal income tax and in the index of affordability (mortgage burden) relate negatively with respect to price, in accordance with theory. Hort (1998) includes the real cost user as one of the long-term determinants of real house prices. She estimates an ECM for Sweden with annual data from 1964 to 1994. The coefficients (depending on the model) range from -0.029 to -0.20 (p. 105). Kuismanen et al. (1999) include user cost in their analysis, although the results show that its effects are "negligible"; the coefficient has the expected sign with is close to 0 (-0.0052, p. 28). McCarthy and Peach (2002) study the reaction of house market to monetary policy in U.S.. They estimate both short-term and long-term price determinants of house prices. In the short-term house prices are unresponsive to user cost of homeownership, with coefficients[24] -0.005 and -0.004. User costs' coefficient of the long-term equation coefficient is -0.029. Although the results have the correct sign, and the magnitudes are seen as reasonable both coefficients are not statistically significant. In Capozza et al. (2004) model, real user cost has negative sing with respect to house prices albeit with a coefficient close to 0 (-0.04, Model 1, p. 17). Riddle (2004) assesses U.S. house market and stock dynamics with a long-term equation by 2SLS. She finds that real user cost of homeownership has a positive sign although with a coefficient near to 0 (0.003), in the same direction than Capozza et al. (2004). Steiner (2010) analyses the house prices and the residential investment in Switzerland from 1975 and 2007. She sets a dynamic model aiming to estimate the dynamics of Swiss house market. She asserts that user cost of homeownership has a significant impact on short-term house prices, with a coefficient of -1.52 (p. 15). Oikarinen (2012)[25] analyses the dynamics between price movements and transactions in Finland, but the user cost of homeownership is

[24] Full sample.
[25] Period from 1988 until 2008.

not relevant in the long run. He believes that the reason they are not entering the long-term equation is due to their nature: they are mean-reverting[26].

Employment (unemployment) is commonly used as proxies or control terms for short-term dynamics as for long-term determinant, and its sign with respect to house price is expected to be positive (negative). For example, Case and Shiller (1990) use employment as a proxy for U.S. metropolitan areas. The expected positive sing is consistent with the theory, and they obtain, depending on the model, coefficients between 0.046 and 0.052 (p. 268); In an equivalent way Abraham and Hendershott (1996) approximate a coefficient of 0.31 for employment (table 2). Kuismanen et al. (1999) use unemployment as an economic variable for the CRH. The sign is negative, as expected: about 1% increase in unemployment decreases house prices in Helsinki area about 0.11% (p. 28). Jacobsen and Naug (2004) find that if unemployment in Norway rises about 4 or 5%, hose prices will fall about 11% in the long run, these results are congruent with the expected effect (p. 37). Wang and Tumbarello (2010) set a ratio of working age population. Theirs results for Australia are consistent with the expected positive sign and with a coefficient 4.38, meaning a 1% increase in the working-age population will raise about 4% the house prices in the long run (p. 11) (quite high as Meese and Wallace (2003) find for Paris or Jacobsen and Naug (2004) for Norway).

Other proxies and house price determinants that appear in the literature are the lending-to-income ratio, loan stock or credit, as a measure of mortgage or credit rationing (see e.g., Meen 1990, Hort 1998[27], Oikarinen 2005, Oikarinen 2009 or Oikarinen 2012). Some literature includes demographic variables or proxies, e.g.,, in Wang and Tumbarello (2010) the ratio of working age population to total population; in Meen (1990) a population proxy caters for demographic changes. In Capozza et al. (2004) population is included to cater for the effect of big cities on rents and prices. A trade proxy as representation for wealth and income since trade shocks are expected to have an impact on income (see e.g., Wang & Tumbarello 2010). Land supply index as in Capozza et al. (2004) to capture the availability of land for developments on house

[26] Hence real user cost does not follow any random walk.
[27] Exogenous variable.

prices. McCarthy and Peach (2002) incorporate the price of land in the residential investment regression (short-term equation). Meese and Wallace (2003) capture the proximity of geographical and infrastructure facilities in Paris with some dummies. Or finally, Steiner (2010) includes a dummy to control price overheating during the 1980s.

Markets adjust slowly, this way, if variables of study are allowed to deviate from their equilibrium levels, it is common to include lagged variables for e.g., when studying the short-term dynamics and their adjustment process. Poterba et al. (1991) found evidence on house price predictability though lagged variables (lagged income and lagged house prices). Although no evidence was found on house prices converging into a long-term level. As explained by Oikarinen (2009), “the slow adjustment, lagged values of the variables are expected to contain informative value with respect to housing price dynamics” (p. 130). The model in Abraham and Hendershott (1996) includes some lagged price variable acting as “bubble-builder”, as prices increase the lagged variable magnifies the process. Shifts in house prices are modelled as the sum of the ECM term or “bubble-burster” and the ”bubble builder” term. Kuismanen et al. (1999) include the lagged income variable. Capozza et al. (2004) find strong serial correlation between the price level and its lagged term (0.33). In Steiner (2010) the price adjustment equation includes the one-period lagged price variable, with strong positive coefficient 0.29 (p. 15).

In Finland Oikarinen (2005) employs a set of several-periods lagged variables (e.g.,, price, number of households, twelve-month Euribor rate)) or as in Oikarinen (2012) where the VECM includes one-period lagged variables of real price, sales volume, real aggregate income and housing stock (p. 50, table 6).

In short-term dynamics, the error correction mechanism, captures the tempo of convergence (negative sign) or divergence towards the long-term equilibrium. For Finland there is consensus that converges at slowly. Takala and Pere (1991) estimate the pace of convergence of prices of 7,4 % per quarter. The ECM sign was as negative, as expected (p. 47). Barot and Takala (1998) estimate that prices converge about a slower pace than Takala and Pere (1991), about 4,4 % per quarter, whereas for Sweden the speed of convergence is slightly higher, about 8-9% per quarter. Oikarinen (2005) finds that prices in Helsinki adjust towards equilibrium at a pace of 9.4% a quarter.

Later, Oikarinen (2009) finds that prices in Helsinki are converging at a slower pace, about 6.4% a quarter, like the results for Finland by Takala and Pere (1991) and Barot and Takala (1998). On the other hand, André and García (2012), although they highlight that convergence is sluggish, they correct more rapid than previous studies about 17% per quarter.

The error corrector term in Drake (1993) shows that the house prices adjust slowly to changes, they are sluggish. Drake states that in the short-term the main drivers of house prices are the lagged values of house completion (supply) and the lagged house prices. DiPasquale and Wheaton (1994) estimate a long-term equation by 2SLS, after a shock the annual rate of price adjustment to equilibrium ranges from 16 to 29% (depending on the model). Abraham and Hendershott (1996) specify a “bubble burster” proxy, which is significant explaining the price shifts in West U.S., although the coefficient is quite small, 0.05, having little explanatory power. Hort (1998) observes that the adjustment towards the long-term equilibrium in Sweden is quite rapid, she estimated coefficient of the error correction term is significant at the 1% (0.836). According to Hort, house prices adjust towards the equilibrium level about 84% every year (p. 111). In Meese and Wallace’s (2003) results suggest a fast adjustment of Parisian dwelling market with a speed of adjustment ranging from 32 to 38 % per month (depending on the model). The second-stage price regression in Capozza et al. (2004) estimates that U.S. house prices converge quickly, about 52% a year.

Some empirical research has supply function as the centre of attention (see e.g., Topel & Rosen 1988, Poterba 1984). Poterba estimates that the supply elasticity is between 0.5 and 2. On the other hand Topel and Rosen estimate a supply-determined model of housing investment of U.S. housing investment with quarterly data series from 1963 until 1983. Results show evidence of strong price-elasticity of investment housing. Short-term supply elasticity (about 1%) seems more moderated than the long-term one (2.76%), but both showing converge within a short span of time, 1 year (p. 735). On the other hand. This supports the underlying idea that supply (construction sector) reacts rapidly to dynamic economic environment (volatility of housing investment) and that supply in the short-term is relatively inelastic (see e.g., Capozza & Li 1994, Dye & Mcmillen 2007, Bulan et al. 2009).

3 RESEARCH METHODS AND DATA

This chapter focuses on research methods and data. In the first place, a description of different methods employed in econometric analysis of house prices. Secondly, an introduction to the estimated empirical long-term model and the estimated short-term model: The former derived from the life-cycle model and the latter from the VECM. Finally, this chapter ends with a comprehensive description of the data gathered for the empirical analysis.

3.1 Research methods

This section begins with defining cointegration and the VECM. It continues covering the Johansen cointegration test. Follows a concise description of how to determine the autoregressive order, followed by an introduction on unit root test, some residual diagnosis techniques, and the structural analysis. Finally, this section covers some extensions of the model and the concept of Granger-causality.

3.1.1 Cointegration

According to Engle and Granger (1987), if the elements of a time series x_t are stationary after taking differences, exists a stationary linear combination $\alpha' x_t$. Then if the time series x_t is cointegrated of order (1,1) with the cointegrated vector α and $\alpha' x_t = 0$, it represents the long-term equilibrium. Cointegration means that equilibrium holds. In some periods x_t will deviate from equilibrium and then $z_t = \alpha' x_t$ is the equilibrium error. If there is cointegration then z_t is stationary as well (because the linear combination is stationary). In words of Enders (2014) that "equilibrium theories involving stationary variables require the existence of a combination of the variables that is stationary" (p. 345).

Formally, if $x_t$is a vector of time series (economic variables): "the components of the vector x_t are said to be co-integrated of order d, b, denoted $x_t \sim CI(d, b)$, if (i) all components of x_t are $I(d)$; (ii) there exists a vector $\alpha (\neq 0)$ so that:

$$z_t = \alpha' x_t \sim I(d, b), b > 0 \tag{27}$$

the vector α is called the co-integrator vector" (Engle and Granger 1987, p. 253).

Thus, a long-term relationship among non-stationary time series exists for instance when they keep the same stochastic trend. This way, their linear combination is stationary, and such time series are cointegrated. Their expansion paths rely on each other and cannot wander away from their long-term equilibrium for ever. If such a long-term relationship can be inferred, the economic implications can be studied as well. Hence, if cointegration among house prices and their fundamentals can be established, their long-term equilibrium exists in a way that any deviation from it is captured through the error. The adjustment process towards the long-term-equilibrium, the short-term dynamics, can be studied from it. However, Enders (2014) warns about cointegration not implying any causality.

If cointegration relations exist in a system of variables, the correct parametrization[28] supporting the analysis is known as ECM or VECM. For this purpose, this thesis estimates a VECM.

To test the existence of one (or multiple) long-term relationships this thesis can make practical and effective use of the Johansen Cointegration test.

3.1.2 Vector Error Correction Model (VECM)

Following Johansen (1988) and Johansen and Juselius (1990), a basic k-th order Vector Autoregressive model (VAR(k)) has the structure[29]:

$$X_t = \Pi_1 X_{t-1} + \cdots + \Pi_k X_{t-k} + \varepsilon_t \ \ for\ t = 1,2, \dots, T \tag{28}$$

[28] An appropriated form of the multivariate system is needed, e.g., whether to include intercept (trend (or not).
[29] Here our model has no deterministic terms, i.e., $\Phi = 0$, and therefore we do not include the vector of centred seasonal dummies D_t and it does not include the µ p-dimensional vector of intercepts. No exogenous variables are included in this simple model (see e.g., Lütkepohl & Krätzig 2004, p. 92-93).

where X_t, is a p-dimensional vector $(x_{1t}x_{2t}$, …, x_{pt})'; each Π_i is a (p x p) coefficient matrix; and ε_t is the i.i.d. p-dimensional Gaussian random variable with zero mean and variance Σ_ε $(\varepsilon_t \sim N(0, \Sigma_\varepsilon)^{30})$. The largest number of lags included in the model is k.

The stability of the process entails that the polynomial defined by the determinant of the autoregressive operator has no root (adapted from Lütkepohl & Krätzig 2004, p. 88, Johansen 1988 and Johansen & Juselius 1990):

$$|A(z)| = (I_k - \Pi_1(z) - \cdots - \Pi_k z^k) \neq 0, \, for \; |z| \leq 1 \qquad (29)$$

where I_k is the identity matrix. From (29), we define the polynomial matrix as $A(z) = I_k - \Pi_1 z_1 - \cdots - \Pi_k z_k$;. If the VAR(k) has no unit root, the determinant of $A(z)$ is different from 0 (z $= 0$).

If (28) has unit root[31] (I(1)), then some or all the variables are integrated and a linear combination with I(0) is feasible. Then $\Pi = A(z)|_{z=1} = I_k - \Pi_1 - \cdots - \Pi_k$ is singular. The X_t sequence is non-stationary and integrated of order 1.

Hence, a more suitable model to accommodate integration relations is the VECM (k-1)[32]. Following Johansen (1988) and Johansen and Juselius (1990) a multivariate model can be generalised, by expressing (28) in first difference form:

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi_k X_{t-1} + \varepsilon_t \qquad (30)$$

Equation (30) is a simple setup of a VECM where ΔX_t is stationary. Matrix $\Gamma_i = -(I - \Pi_1 - \cdots - \Pi_i)$ with $i = 1, \ldots, k-1$.

[30] $E(u_t u_t') = \Sigma_u$: positive definite covariance matrix.
[31] $det = (I_k - \Pi_1 z - \cdots - \Pi_k z^k) = 0$, for $|z| = 1$.
[32] Considering that error correction term is at lag t-1.

The terms $\Gamma_j (j = 1, \dots, k-1)$ from (30) ($\Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1}$) represent the short-term dynamics of the models. As in (28), the term $\Pi_k X_{t-1}$ is I(0)[33] and comprises the cointegration relationships. It holds information about the long-term interaction among variables and therefore represents the long-term model, in words of Oikarinen (2007) "the error correction term of the model" (p. 67). Rewriting Π:

$$\Pi = \alpha\beta' \tag{31}$$

where α and β are matrices with dimension p x r, and $\mathrm{Rank}(\alpha) = \mathrm{Rank}(\beta) = \mathrm{r}$. Matrix β is the cointegration matrix of the system; $\Pi_k X_{t-1}$ based on (28) can be rewritten as $\alpha\beta' X_{t-1}$, even though X_{t-1} sequence is non-stationary, the linear combination $\beta' X_t$is stationary and contains the cointegration relationship. And matrix α comprises the speed of adjustment parameters. Low coefficients reveal a slow correction process towards the estimated equilibrium state whereas high values indicate a rapid adjustment (Johansen & Juselius, 1990). According to Johansen and Juselius (1990), expression (28) can be converted in a VECM if Π rank, the number of cointegration relationships, matches with the number of independent cointegrating vectors. Thus:

- $\mathrm{Rank}(\Pi) = \mathrm{r} = \mathrm{p}$, then matrix has full rank. Hence X_t is stationary, no need to Estimate any VECM a Vector Autoregression model (VAR) is enough.
- $\mathrm{Rank}(\Pi) = 0$, then there is no cointegration. VECM does not exist.
- $0 < \mathrm{Rank}(\Pi) = r < p$, then there exist matrices α and β from (30) such that a VECM can be estimated.

3.1.3 Specify the cointegrating rank of the VECM.

Johansen Cointegration test[34] is a procedure for testing cointegration and to determine the correct parametrization of the VECM.

---

[33] This term is the only one consisting of I(1) variables (X_{t-1}).
[34] All the variables of the model have to be integrated of the same order, I(1).

The Johansen procedure stands on the maximum likelihood estimator, $\hat{\beta}$, when running tests for the relationship given in (31) (Johansen, 1988), the null hypothesis is:

$$H_0: rank(\Pi) \leq r \;\; or \; \Pi = \alpha\beta'$$

In practice, two test statistics are used (from Johansen & Juselius 1990):

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{p} ln\left(1 - \hat{\lambda}_i\right) \qquad (32)$$

$$\lambda_{max}(r, r+1) = -T\, ln\left(1 - \hat{\lambda}_{r+1}\right) \qquad (33)$$

where $\hat{\lambda}_i$ are the estimated values of the eigenvalues obtained from estimated matrix Π; and T is the number of observations.

The likelihood ratio test (32) tests for the null hypothesis, $H_0: rank(\Pi) \leq r$, and when $\lambda_i = 0 \; then \;\; \lambda_{trace} = 0$. This test statistics (33) tests for the null hypothesis, $H_0: rank(\Pi) = r$ against an alternative hypothesis of r + 1 cointegrating vectors, when the estimated value of the eigenvalue,λ_{r+1}, is close to zero then λ_{max} is small (Enders, 2014, p. 378). If the values of the test statistics are higher than the critical values null hypothesis are rejected and therefore cointegration. Asymptotic distributions of t-statistics and estimators are conditional on model formulation, hence test statistics are not always given by a χ^2 distribution, i.e., inferences on vectors α and β can be conducted based on χ^2 under linear restrictions (Johansen & Juselius 1990, p. 169).[35]

3.1.4 Determining the autoregressive order

For the purpose to decide the correct model specification some information criteria are employed. According to Lütkepohl and Krätzig (2004) a well fit model has to include all the terms of interest for the relationship under study. Hence, choosing the appropriate lag length (k) will increase the performance of the model. A lag length k

[35] Statistical software integrates them.

too large consumes degrees of freedom. On the other hand, if lag length is too small, the model may be lame.

The information criteria used in this econometric analysis are the Akaike Information criterion (AIC), the Schwarz information criterion (SBC)[36] and Hannan-Quinn criterion (HQ):

From (Lütkepohl & Krätzig 2004[37]):

$$AIC(m) = log\,det(\widetilde{\textstyle\sum_u(m)}) + \frac{2}{T} mp^2 \tag{34}$$

$$SBC(m) = log\,det(\widetilde{\textstyle\sum_u(m)}) + \frac{log\,T}{T} mp^2 \tag{35}$$

$$HQ(m) = log\,det(\widetilde{\textstyle\sum_u(m)}) + \frac{2log\,log\,T}{T} mp^2 \tag{36}$$

where m is the order of the model such that $m = 0, \dots, k_{max}$, opting for estimators of order k that minimise the chosen criterion; $\widetilde{\sum_u(m)}$)represents the variance-covariance matrix estimator of order m. Criterion (34) stands for Akaike test, (35) for Schwarz test and (36) for Hannan-Quinn test.

Each information criterion will determine a lag length. We shall select the information criterion that minimizes m and guarantees enough lags to capture the dynamics of the model. In practice, criteria with low values are preferred[38].

3.1.5 Augmented Dickey-Fuller test[39]

The augmented Dickey-Fuller test (ADF) evaluates stationarity or non-stationarity of time series. Estimating a VECM requires that time series be non-stationary in levels.

---

[36] Known as well as Bayesian information criterion.
[37] Adapting notation $K \rightarrow p$.
[38] Although residual autocorrelation may modify lag length choice.
[39] Other tests: Phillips-Perron (PP) tests or ADF-GLS test.

The next step, once the lag length is determined, is to test for unit root (I(1)). A stochastic process x_t is stationary if[40]:

$$E(x_t) = \mu_x \; for \; all \; t \in T \tag{37}$$

$$E[(x_t - \mu_x)(x_{t-h} - \mu_x)] = \gamma_h \tag{38}$$
$$for \; all \; t \in T \; and \; all \; integers \; h \; such \; that \; t - h \in T$$

Thus, (37) implies that all members of a random process have equal and constant mean. This way, the stationary random process generates time series that wander around a constant mean, without trend. According to Engle and Granger (1987) a shock has only temporary effect on it.

Expression (38) guarantees that the variances and covariances are time invariant (Lütkepohl & Krätzig 2004, p. 11).

Although, stationarity is an unusual quality of economics time series, Lütkepohl and Krätzig (2004) tell that it is possible to get stationary-looking time series by transforming them. According to Lütkepohl and Krätzig (2004) trend-stationary processes and asymptotically stationary processes can be made stationary by subtracting trend or by modifying some initial values (p. 12). Let be x_t an autoregressive process of order k (AR(k)). This AR(k) is integrated when $\alpha(1) = 1 - \alpha_1 - \cdots - \alpha_k = 0$

$$x_t = \alpha_1 x_{t-1} + \cdots + \alpha_k x_{t-k} + \varepsilon_t \tag{39}$$

where α_i are coefficients and ε_t is an unobservable white noise process with mean 0 and constant variance ($E(\varepsilon_t^2) = \sigma_\varepsilon^2$). Transforming equation (39) with the subtraction of x_{t-1} on both sides:

[40] Based on Lütkepohl and Krätzig (2004) and modified: $y_t \rightarrow x_t$; $p \rightarrow k$; $u_t \rightarrow \varepsilon_t$.

$$\Delta x_t = \phi\, x_{t-1} + \sum_{j=1}^{k-1} \alpha_j^* \,\Delta\, x_{t-j} + \varepsilon_t \tag{40}$$

where $\phi = -\,\alpha(1)$ and $\alpha_j^* = -(\alpha_{j+1} + \cdots + \alpha_k)$ (based on Lütkepohl & Krätzig 2004, p. 54)

The null hypothesis $H_0\colon \phi = 0$ (time series is non-stationary) is tested against the alternative hypothesis $H_s\colon \phi < 0$ (time series is stationary). Estimating (40) with an Ordinary Least Squares (OLS) gives the t-statistics of parameter ϕ. The inferences are based on the t-statistics of parameter ϕ. The t-statistics of estimated ϕ is found on some tables in tables in Dickey and Fuller (1981).

More specifically, if ADF's null hypothesis, H_0, is not rejected, the time series are stationary in levels, then cointegration is not possible.

3.1.6 Residual analysis

A set of tools are commonly used to determine whether the VECM provides an accurate representation of the data-generating process (DGP). Most of them are studying the residuals (Lütkepohl & Krätzig 2004).

The assumption is that residuals are i.i.d.[41] and $\varepsilon_t \sim N(0, \Sigma)$ and $Corr[\varepsilon_t, \varepsilon_{t-j}] = 0\ \forall j \neq 0$.

To check for residual autocorrelation, a non-formal approach is based on visual inspection of the correlograms when using statistical software. A more formal test for residual autocorrelation is the Lagrange multiplier (LM)[42] test.

[41] Independent and identically distributed.
[42] An alternative test for residual autocorrelation is the Portmanteau test.

The LM test for h-th order residual autocorrelation assumes a model (From Lütkepohl & Krätzig 2005[43]):

$$\varepsilon_t = B_1^* \varepsilon_{t-1} + \cdots + B_h^* \varepsilon_{t-h} + error_t \tag{41}$$

Its null hypothesis, H_0, is no autocorrelation: H_0: $B_1^* = \cdots = B_h^* = 0$ vs H_1: $B_1^* \neq 0$ $or \ldots or$ $B_h^* \neq 0$.

If the LM tests rejects the null hypothesis (H_0), then the model is incorrectly specified, and more lags have to be included in the VECM.

The LM statistic is:

$$LM_h = T[p - tr(\widetilde{\Sigma_e}\widetilde{\Sigma_R^{-1}})] \tag{42}$$

This test is preferred if testing short lag autocorrelation. The statistic (42), under standard assumption has an asymptotic $\chi^2(hp^2)$ distribution.

Another step is to test for VECM residual normality, whether the residuals are symmetrically distributed[44] or not. A multivariate Jarque-Bera normality test[45] assesses residual normality. Under null hypothesis, H_0, residuals follow a normal distribution; it is tested against non-normal distribution. Jarque-Bera statistic has a $\chi^2(2p)$[46] limiting distribution (Lütkepohl & Krätzig 2004, p. 130).

A third step is to detect autoregressive conditional heteroskedasticity (ARCH effects). The basic assumption is to believe that residual series are homoscedastic. The ARCH–LM test is performed to test ARCH effects on the residuals. The null hypothesis

[43] Modified original: $K \to p$; $u_t \to \varepsilon_t$.
[44] And hence, kurtosis and skewness are zero.
[45] Known as well as Lomnicki–Jarque-Bera test for nonnormality (Lütkepohl & Krätzig, 2004).
[46] Modified from original: $K \to p$.

assumption, H_0, is that residuals series are homoscedastic. The results from the ARCH-LM statistic are compared with critical values from a χ^2 distribution.

However, according to Lütkepohl and Krätzig (2004) “remaining ARCH in the residual series may not be a big problem” (p. 131). In the same vein, there is some evidence that moderate ARCH effects do not affect the robustness of the cointegration rank tests. (Cavaliere, Rahbek & Taylor 2010 and Oikarinen 2007)

3.1.7 Stability analysis

According to Oikarinen (2007) structural breaks may occur over time and affect cointegration interactions, and hence may alter the long-term coefficients of our model. Some external and unpredicted event may cause a change in the pattern of the data. (e.g.,, Hansen (2003) studies the effects of the change in the term structure of interest rates in U.S. from 1979-1982; Oikarinen (2009) test for possible structural breaks in Finnish housing market due to the European Monetary Union; Meen (1990) evaluates the removal of mortgage market constraints in U.K. (end of mortgage rationing happened during 1980-1981)).

Chow (1960) say that “no economic rationale in assuming that two relationships are completely the same. It may be more reasonable to suppose that only parts of the relationships are identical in two periods” (p. 591). Thus, a structural break occurs in a period T_B if a model is estimated from a full sample of T observations and $T_1 < T_B$ and $T_2 \leq T - T_B$. T_1 represents the first observation and T_2 the last observation. (Lütkepohl & Krätzig 2005).

According to Lütkepohl and Krätzig (2005) different statistics can be computes, the break-point statistic (BP), the same-split (SS) test and the forecast test (CF) is:

$$\lambda_{BP} = (T_1 + T_2)\, log\, det\, \tilde{\Sigma}_{(1,2)} - T_1\, log\, det\, \tilde{\Sigma}_{(1)} - T_2\, log\, det\, \tilde{\Sigma}_{(2)} \qquad (43)$$

$$\lambda_{SS} = (T_1 + T_2)[log|\tilde{\Sigma}_{1,2}| - log|\{(T_1 + T_2)^{-1}(T_1\tilde{\Sigma}_{(1)} + T_2\tilde{\Sigma}_{(2)})\}|] \qquad (44)$$

Expression (43) represents the break-point statistic and (44) the same-split (SS). The results from (43) and (44) are compared with critical values from a χ^2 distribution, where the difference among the total of parameters estimated in the full sample model and in the initial and the final subperiods determine the degrees of freedom. The null hypothesis, H_0, tests whether the estimated number of parameters is constant or not. If results of (43) and (44) are large, then H_0 is rejected, and there is structural break. Test (44) assumes constant white noise covariance matrix (Lütkepohl & Krätzig 2005, p. 28).

Finally, in the forecast test (CF)[47] for a break point at T_B:

$$\lambda_{CF} = \frac{1-(1-R_r^2)^{1/s}}{(1- R_r^2} \times \frac{N_s-q}{pk^*} \approx F(pk^*, Ns-q) \tag{45}$$

here k^* is the quantity of predicted periods ($k^* = T - T_1$); k_1 is the quantity of regressors in the time-invariant model; $N = T - k_1 - k^* - (p - k^* + 1)/2$; $s = (\frac{p^2 k^{*2} - 4}{p^2 + k^{*2} - 5})^{1/2}$; $q = \frac{pk^*}{2} - 1$; $R_r^2 = 1 - (\frac{T_1}{T})^p \left|\tilde{\Sigma}_{(1)}\right| (\left|\tilde{\Sigma}_{(\mathrm{u})}\right|)^{-1}$. The CF checks against the alternative that all coefficients including the white noise covariance matrix may fluctuate. The H_0 is rejected for large values of the $F(\cdot)$ (Lütkepohl & Krätzig 2005, p. 29).

3.1.8 Extensions of the model[48]

A more generalized form of VECM can be obtained by adding some deterministic terms, such as some intercept, a linear trend term or seasonal dummies (Lütkepohl & Krätzig 2004), by modifying model (30):

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi_k X_{t-1} + \mu_t + \Phi D_t + \varepsilon_t \tag{46}$$

[47] Adapted from Lütkepohl and Krätzig (2004) to the notation of this thesis: $K \to p$.
[48] From Johansen and Juselius (1990), Oikarinen (2007), Lütkepohl and Krätzig (2004) and SAS (2014).

where $\mu_t = \mu_0 + \mu_1 t$ is the deterministic term vector; D_t, some (s-1)-dimensional vector of centred seasonal dummies and Φ is a (p x (s-1))-dimensional coefficient matrix.

The inclusion of deterministic terms such as a linear term trend, a constant or dummy variable (seasonal) may be required to avoid misrepresenting the DGP (Lütkepohl & Krätzig 2004). Thus:

$$x_t = \mu_0 + \mu_1 t + y_t{}^{49} \tag{47}$$

where y_t is a stochastic process (VECM or VAR interpretation); $\mu_1 t$ is a linear trend and μ_0 an intercept term.

If $v_0{}^* = -\Pi\mu_0 + (\sum_{i=1}^{k} iA_j)\mu_1$, $\mu_1 = 0$ and $\Pi^* = [\Pi\colon v_0{}^*]$ is a (p x (p +1)) matrix and relationship (31), we can study the deterministic trend of the model in three different cases:

Case 1) If $\mu = \alpha\beta_0'$ when $\mu = 0$ in (46), there are no deterministic terms in the model and becomes model (30). There are no deterministic terms in the model. There are neither deterministic trends in data nor trends or intercepts in the error correction term. (30) Is the most restrictive model. And $\Pi^* = \alpha\beta^{*\prime}$.

Case 2) When $\mu \neq 0$, a less restrictive version, as in (48), there is no separate drift in the VECM, but intercepts enters only though the error correction term (SAS 2014), and therefore no linear trend in data. Thus:

$$\Delta x_t = \Pi^* \binom{x_{t-1}}{1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t \tag{48}$$

[49] Notation modified: $y_t \rightarrow x_t$, $K \rightarrow$ p and $j \rightarrow$ i.

Case 3) As in (49), when $\mu_1 = 0$, the non-stationary process X_t has a separate drift in the short-term of the model (v_0), the model does not exhibit different linear trend in the error correction term:

$$\Delta x_t = \Pi_k x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + v_0 + \Phi D_t + \varepsilon_t \tag{49}$$

A second group of cases include a linear deterministic trend, $\mu_1 \neq 0$. In these cases $v_0 = -\Pi\mu_0 + (\sum_{i=1}^{k} iA_j)\mu_1$.

Case 4) When $\mu_1 \neq 0$, there is a separate drift in the VECM, but linear trend only enters through the error correction term (long-term part of the equation):

$$\Delta x_t = v_0 + \Pi x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t \tag{50}$$

In (50), $\mu_1 \beta' = 0$ and this implies that $r < p$.

Finally, case 5, when the model has a separate linear trend in Δx_t. This is known as the least restrictive model:

$$\Delta x_t = \nu + \Pi^+ \binom{x_{t-1}}{t} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t \tag{51}$$

where the (p x (p + 1)) matrix of rank r, $\Pi^+ = \alpha[\beta' : \eta]$; $\eta = -\beta' \mu_1$ and $\nu = -\Pi\mu_0 + (I_k - \Gamma_1 - \cdots - \Gamma_{k-1})\mu_1$. These assumptions indicate that variables and the long-term model have some deterministic trend (Lütkepohl &Krätzig 2004, p. 114).

Another generalization of the model considers including exogenous variables in the model (46). Apart from the deterministic part of the model, in practice a stochastic part can be added to complete the model, although according to Lütkepohl and Krätzig (2004), this has to be done cautiously to accomplish with the exogeneity requirements: that none of the exogenous variables are reliant on the dependent variable of the model (X_t).

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Pi_k X_{t-1} + BZ_t + \mu + \Phi D_t + \varepsilon_t \tag{52}$$

where B is a parameter matrix and Z_t is a p-dimension vector of exogenous variables.

According to Lütkepohl and Krätzig (2004): “a set of variables z_t is said to be weakly exogenous for a parameter vector of interest, for instance θ, if estimating θ within a conditional model (conditional on z_t) does not entail a loss of information relative to estimating the vector in a full model that does not condition on z_t” (p. 92).

According to Enders (2014) Z_t is weakly exogenous if it does none of the error correction and does not respond to divergences from long-term equilibrium of $\mathrm{X_t}$. Hence, its speed adjustment parameter is 0. Moreover, the parameters that determine Z_t have to be different from those determining $\mathrm{X_t}$, for instance, lagged value of $\mathrm{X_t}$ cannot determine Z_t. Weakly exogeneity is needed for estimation.

3.1.9 Granger-causality

According to Lütkepohl and Krätzig (2004) a variable y_{2t} has causal effects on another variable y_{1t} , if “the former helps to improve the forecast of the latter” (p. 144). If including lagged values of y_{2t} when forecasting y_{1t} helps to improve its forecast, then Granger-causality may not exist if:

$$y_{1t+h|\Omega_t} = y_{1t+h|\Omega_t \backslash \{y_{2,s|s\leq t}\}}, where\ h = 1,2,\ldots. \tag{53}$$

where $y_{1t+h|\Omega_t}$ represents the optimal h-step forecast of y_{1t} at the origin t founded on the essential information in the universe Ω_t.

A bivariate VECM can be written as[50]:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \alpha\beta' \begin{bmatrix} y_{1,t-1} \\ y_{2,t-2} \end{bmatrix} + \sum_{i=1}^{k-1} \begin{bmatrix} \gamma_{11,i} & \gamma_{12,i} \\ \gamma_{21,i} & \gamma_{22,i} \end{bmatrix} \begin{bmatrix} \Delta y_{1,t-i} \\ \Delta y_{2,t-i} \end{bmatrix} + \varepsilon_t \tag{54}$$

[50] Notation modified: $p \rightarrow$ k and $u_t \rightarrow \varepsilon_t$.

If $\alpha\beta' = \begin{bmatrix} \alpha_1\beta_1 & \alpha_1\beta_2 \\ \alpha_2\beta_1 & \alpha_2\beta_2 \end{bmatrix}$. Then (53) is comparable to test the null hypothesis (H_0) that lagged values of y_{2t}, as if $\gamma_{12,\mathrm{i}} = 0, where\ i = 1, \ldots, k-1$ and that $\alpha_1\beta_2 = 0$. Wald test statistic assesses the H_0.

3.2 Theoretical model[51]

On behalf of analysing the short-term dynamics and the adjustment towards an estimated long-term equilibrium we estimate one VECM for the CRH, using the Johansen procedure. This section introduces the models to estimate in the data analysis section (section 4).

In the same vein than Meen (1990) and Oikarinen (2009), we take as starting point the life-cycle model. To better compare our results with those obtained in Finland we are going to rearrange and specify equation (25) and equation (26) in the same way as Oikarinen (2009):

$$U_t = (1 - T_t)i_t + \gamma - \pi + \delta - E(P_{t+1} - P_t) + {}^{\lambda_t}/_{\mu_c} \qquad (55)$$

$$P_t = {}^{R_t}/_{U_t} \qquad (56)$$

where U_t is the real user cost of homeownership; $(1 - T_t)i_t$ is the after-tax market interest rate; γ is a risk premium to cater the risk of owning with respect to renting, we added this component following Oikarinen notation, it is assumed time-invariant; π is the inflation rate; δ is the depreciation rate (preservation and repairs); $E(P_{t+1} - P_t)$ is expected real capital gain of owning housing; ${}^{\lambda_t}/_{\mu_c}$ represents a shadow price of mortgage rationing (λ_t) divided by the marginal utility of consumption (μ_c); P_t[52] is the real house price level and R_t is the real imputed rental price of housing services.

[51] Specially Oikarinen (2009) provides a baseline for this section.
[52] Real house price and house price are used interchangeably.

Equation (55) represents how the real user cost (of ownership) is determined and the equation (56)[53] represents the "arbitrage condition" of the housing market. R_t is not directly observable and has to be proxied. We assume as in Oikarinen (2009) that it is determined by the real aggregate income of the area of interest Y_t, some demographic variables D_t, and the housing stock H_t:

$$\underset{}{R_t = f(\underset{+}{Y_t}, \underset{+}{D_t}, \underset{-}{H_t})^{54}} \tag{57}$$

Therefore, for empirical purposes R_t is replaced by its determinants in the model. Computing correctly U_t entails difficulties and some adjustments are required, in the next section we discuss about how to overcome this issue. The house prices are determined by supply and demand determinants. All the determinants in this thesis driving house prices are fundamental demand factors but construction costs (CC_t) and housing stock (H_t), which are fundamental supply factors.

Real house prices are driven by the following fundamentals:

$$P_t = f(\underset{+}{Y_t}, \underset{+}{L_t}, \underset{-}{U_t}, \underset{-}{H_t}, \underset{+}{CC_t}) \tag{58}$$

Real house prices are a function of the Y_t[55], L_t, the (outstanding) loan stock variable[56], U_t, H_t and CC_t. Hence, the long-term model (log-linearising) to estimate is:[57]:

$$P_t = \beta_1 Y_t + \beta_2 L_t + \beta_3 U_t + \beta_4 H_t + \beta_5 CC_t + \varepsilon_t \tag{59}$$

[53] This condition helps to evaluate whether house prices are above their level (or below).

[54] Meen (1990) includes household wealth, but Oikarinen (2009) remains us that there is not enough data regarding household wealth and is left outside the model.

[55] From now onwards aggregate income caters for real per capita income and population.

[56] This is not included in the life-cycle model, according to Oikarinen (2009) is a good proxy for λ_t/μ_c.

[57] Long-term model is similar to the one in Oikarinen (2012), although construction costs (CC_t) are included. All variables but U_t are in logarithms.

where, β_i, for $i = 1, \dots, 5$, represent the coefficients of the variables and ε_t is a white noise error term. We take natural logs of all variables but U_t. The matrix β of coefficients has to be estimated to find the long-term relationships.

As Oikarinen (2009) retells, the cointegration relationship of (59) if assumed, imposing that the house prices cannot drift apart endlessly from their long-term equilibrium level.

In the short run, house prices are expected to deviate from their equilibrium as a result of a shock. Nonetheless the adjustment back towards the long-term equilibrium level is achieved with sluggishness (see e.g., Drake 1993, DiPasquale & Wheaton 1994, Barot & Takala 1998, Malpezzi 1999, Oikarinen 2005 and Oikarinen 2009). A similar VECM model as the one presented in (46)[58] is estimated:

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} + \alpha' \epsilon_{t-1} + \mu + \Phi D_t + \varepsilon_t{}^{59} \qquad (60)$$

where is a X_t is a six-dimensional vector consisting of P_t, Y_t,L_t, U_t, $H_t$and CC_t; Γ_{k-1} is a six times six matrix of coefficients for the stochastic differenced variables at lag i. In our model k-1 lags; μ is a six-dimensional deterministic term vector; D_t is a three-dimensional vector of seasonal dummies (s-1, since we are using quarterly data, s=4;S_2, S_3 and S_4); Φ represents a six times three coefficient matrix and ε_t is a six-dimension vector of white noise residuals. (Oikarinen 2009)

As in Oikarinen (2009) the long-term part of the model (the error correction term) is $\alpha' \epsilon_{t-1}$, where vector α' contains the speed of adjustment parameters towards the equilibrium and the one-term lagged deviation of house prices from their long-term level are embodied in the error correction term ϵ_{t-1}, which can be represented as:

$$\epsilon_{t-1} = P_{t-1} - \beta_1^* Y_{t-1} - \beta_2^* L_{t-1} - \beta_3^* U_{t-1} - \beta_4^* H_{t-1} - \beta_5^* CC_{t-1} \qquad (61)$$

[58] $\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi_k X_{t-1} + \mu + \Phi D_t + \varepsilon_t$
[59] Similar to Oikarinen (2009).

Fundamental variables are, a priori included in both, the long-term and the short-term model as endogenous variables following Oikarinen (2009) indication, which suggests that there is enough empirical support in the case of Finland for this to happen. In section 4, we conduct the necessary examination and tests to verify and determine the correct shape of our econometric model. It includes a more detailed and precise description on the need for exclusion or not, if necessary, of some of the fundamentals in the econometric model; as well as the appropriated tests for weakly exogeneity; or the inclusion of any deterministic terms in the model, when testing the order of integration.

3.3 Data[60]

This section includes the description and justification of the data being used (contents, sources, quality, adjustments made, etc…). The data comprises the period 1995Q1-2019Q4 (quarterly and quarterly adjusted). We use 2010=100 as base index in the cases nominal data has to be deflated to real values. We take natural logs of all variables but U_t, since might have negative values.

3.3.1 Aggregate income (Y_t)

To conduct time series analysis on housing prices, the first data we use is an index created and provided by professor Oikarinen. He computed a quarterly index that describes the development of after-tax income at the CRH level. According to Oikarinen (2009). The sources of this data are Tilastokeskus (annual disposable income and the population of the CRH area) and professor Oikarinen[61] himself. Previously, the same index has been employed in Oikarinen (2009) or Oikarinen (2014). The nominal index is deflated (2010=100). The same index is employed to conduct CRH analyses.

[60] Figures in this section are in natural logarithms.
[61] For more information regarding its estimation see Oikarinen (2014, p. 1690).

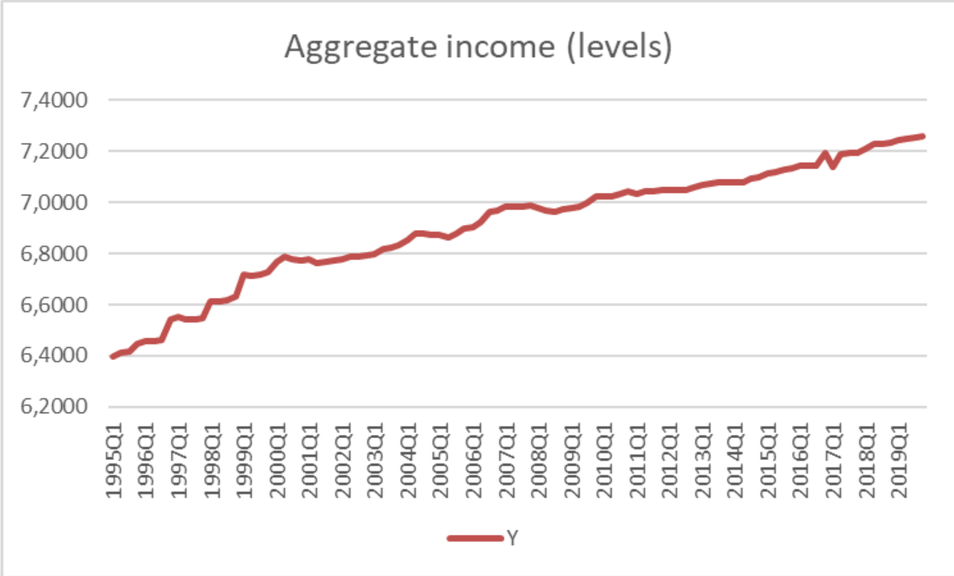

Figure 7. Evolution of aggregate income (levels) in CRH 1995Q1-2019Q4. (Source Tilastokeskus and professor Oikarinen)

Demographic data is published by Tilastokeskus, on a regional and monthly basis. We use this variable along with net disposable income in order to capture demographic demand trends.

3.3.2 Real house prices (P_t)

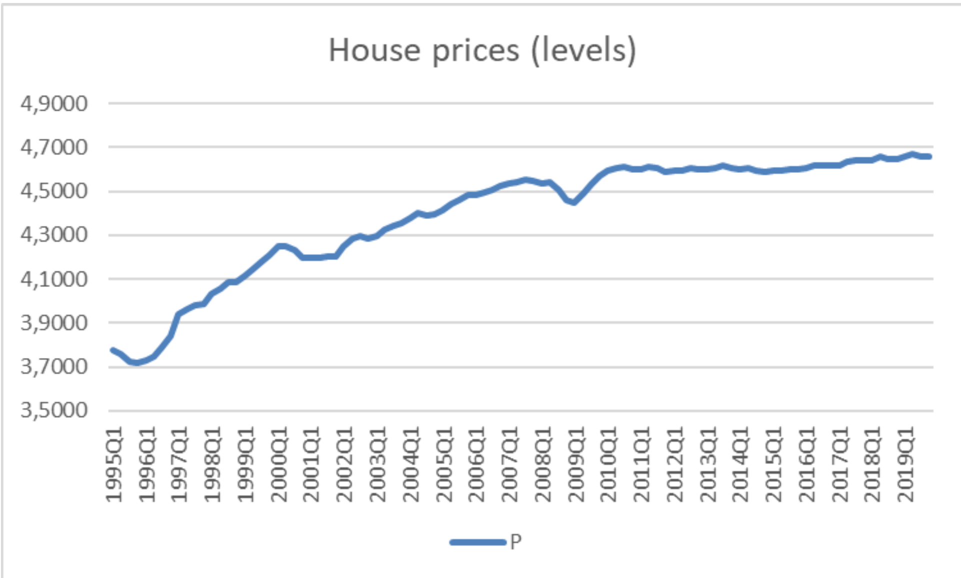

Figure 8. Evolution of house prices (levels) in CRH 1995Q1-2019Q4. (Source Tilastokeskus)

Consists of data on the real price index for CRH published by Tilastokeskus (Statistics Finland). We use the raw data series “Real Price Indices of old dwellings in housing companies”, with the nominal price index already deflated by the Consumer Price

Index (CPI) of the corresponding period and year. A single real price index is constructed by chaining 2 indices using the base index (2010=100). The Price index series cover from 1995Q1 until 2019Q4 using quarterly data. These data series describe the sequence of old dwelling's prices in CRH for the "total building types" and the "total number of rooms". The price index is based on the price information collected by Vero for tax purposes. The real price index serves as a perfect tool to compare the present average changes in prices of dwellings with the past. In order to calculate the price index, Tilastokeskus uses some hedonic method, separating the changes in prices due to the market and those due to the asset own characteristics.

3.3.3 Real construction costs (CC_t)

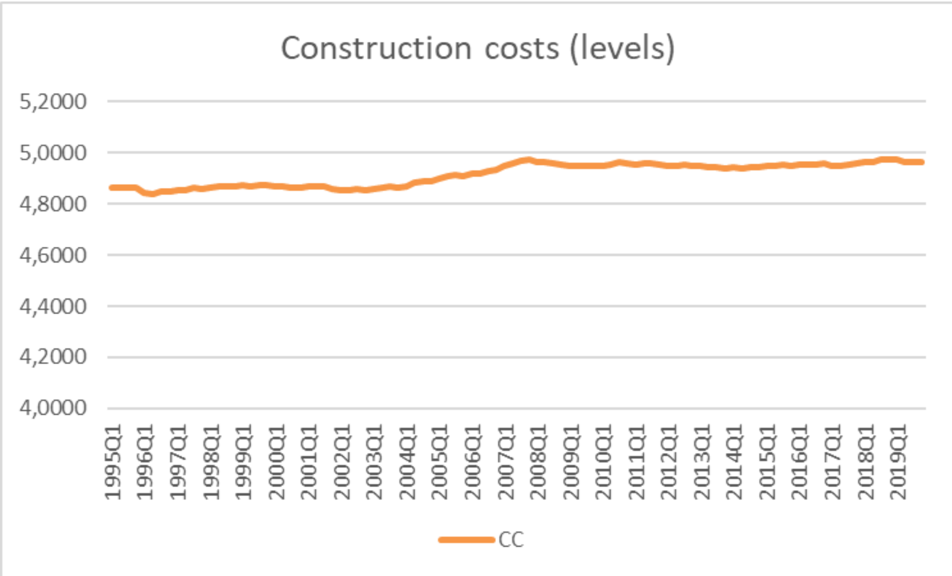

Figure 9. Evolution of construction costs (levels) in Finland 1995Q1-2019Q4. (Source Tilastokeskus)

Data is published by Tilastokeskus as monthly data on the "Building Cost Index by type of building" from 1995Q1 to 2019Q4. This data series describes the evolution of construction costs at national level, monitoring the change in the prices of the production factors used during the construction process, such as materials, wages and salaries and any other input regarding construction. The data selected covers the subsection "block of flats". Two building cost index series have been chained using changes in the latest index (2010=100). Unfortunately, as in Hort (1998), this cost index takes not into account any discrepancies in terms of the quality of dwellings. We have converted monthly data into quarterly data.

Oikarinen (2009) says that construction costs have not been included either on the long-term or short-term model, since have shown a constant trend during the period covered by our empirical research, and he believes that his effect on house prices can be ignored. Yet, we decided to include them. Construction costs are estimated at national level, since there is no data disclosed at an inferior level.

3.3.4 Housing stock (H_t)

This variable captures the volume of new construction entering the market. Tilastokeskus publishes monthly data on “Building and dwelling production”. Unfortunately, the published data is at national level. Monthly data is converted into quarterly data. We have selected the number of dwellings for all kind of residential buildings[62]. The construction stage includes the total of building completions.

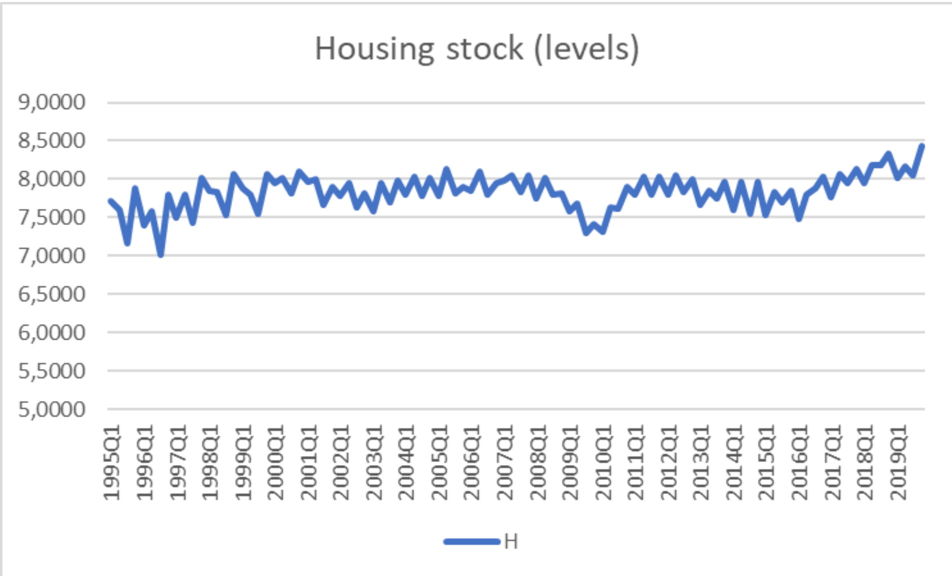

Figure 10. Evolution of housing stock (levels) in Finland (unadjusted) 1995Q1-2019Q4. (Source Tilastokeskus)

Capital regions tend to concentrate a great share of the economic activity, jobs, and investments of a country. Therefore, the influence of the Helsinki metropolitan region exerted on Finland cannot be disregarded, in words of Oikarinen (2007) “national

[62] Detached and semi-detached, blocks of flats, residential buildings for communities and dwellings for special groups.

housing price figures are dominated by housing in the HMA[63]" (p. 217). It is however reasonably to assume that CRH construction development resembles the one described by national data.

3.3.5 Real user cost of homeownership (U_t)

The correct estimation of U_t (measured in percent) entails difficulties, for this reason some adjustments are required to equation (55). We have simplified the estimation of the real user cost of ownership, in such a way that it is equivalent to the real after-tax mortgage rate proposed by Oikarinen (2007). Hence our simplified version of equation (55) is:

$$U_t = (1 - T_t)i_t - \pi + \delta \qquad (62)$$

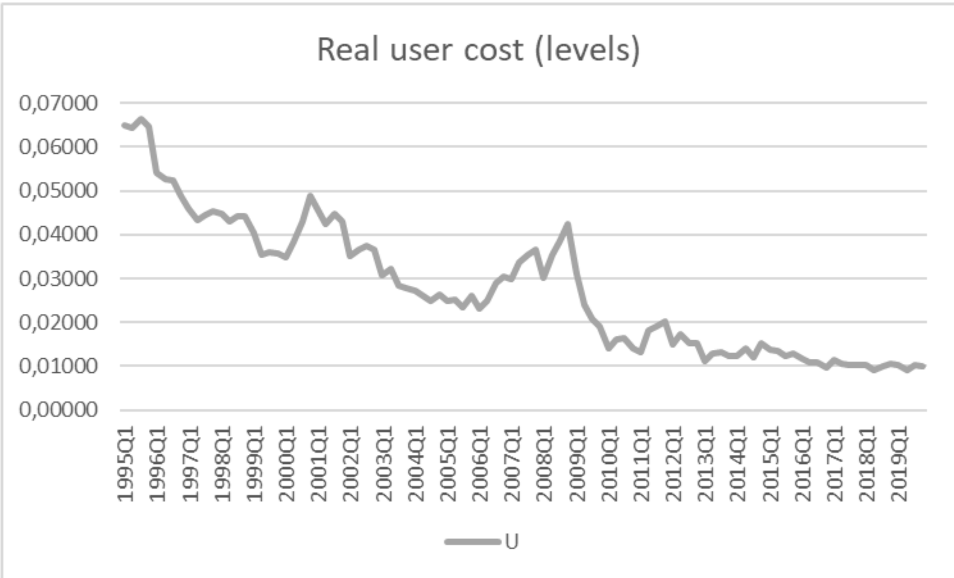

Figure 11. Evolution of real user cost (levels) in CRH 1995Q1-2019Q4. (Different sources)

In the same vein than Hort (1998) or Oikarinen (2005) we do not include depreciation and maintenance rate as part of δ, thus, it is tacitly assumed constant and not considered in our empirical analysis. In Finland, property tax (kiinteistövero) is not deductible (Oikarinen 2007), thus its value should be included as part of δ (Oikarinen 2009). Data series from tax rates (vakituisen asuinrakennuksen kiinteistöveroprosentti)

[63] HMA=CRH

are constructed from different sources: Kaleva, Niemi and Törnroos (n.d., p. 46), Helin (2007, p. 12), Verohallinto (https://www.vero.fi/)[64] and Vantaan Kaupungin tilinpäätos 2009. Property tax rates are:

- 1995-1999: 0.2%
- 2000-2009: 0.22% (Helsinki); 2003: 0.32% (Vantaa); 2004-2007: 0.42% (Vantaa); 2008-2009: 0.3% (Vantaa); 2009: 0.32% (Espoo and Kauniainen).
- 2010-2014: 0.32%
- 2015-2016: 0.37%
- 2017-2019: 0.41%

We use as reference Helsinki. Poterba (1984) says that tax deductibility of loan payments makes homeownership more attractive. In Finland selling the permanent home is exempt from taxation[65].

According to Oikarinen (2009) taxpayers deduct their interest spending multiplied by the capital income tax rate (ever since the 1993 tax reform). To approximate the nominal after-tax market interest rate, $(1 - T_t)i_t$: we use as i_t, the monthly average interest rate of housing loan published by the Suomen Pankki (Central Bank of Finland, https://www.suomenpankki.fi/)[66]. The data is published monthly since June 1989 on the average lending rates of financial institutions in Finland. We have converted monthly data into quarterly data. The capital income tax rates (T_t) is constructed merging diverse sources: Acts Amending the Income Tax, Verohallinto (2015) and Pöörsisäätiö (2020). Since 1995 has fluctuated from 25% to 30%[67]:

---

[64] https://www.vero.fi/tietoa-verohallinnosta/tilastot/kiinteistoverotilastoj/
[65] Selling qualifies for tax exemption if the owner has permanently lived for at least two years or if the proprietor has owned the dwelling for at least two years.
[66] From:https://www.suomenpankki.fi/en/Statistics/mfi-balance-sheet/tables/rati-taulukot-en/talletusten_ja_lainojen_korot_en/
[67] For capital incomes exceeding a threshold (nowadays 34,000 euros) applies a second income tax rate (at the moment is 34%) but it is not considered.

- 1995: 25%
- 1996–1999: 28%
- 2000–2004: 29%
- 2005–2011: 28%
- 2012–2019: 30%

The real after-tax mortgage rate is calculated by detracting the inflation (π) from the nominal after-tax mortgage rate ($(1 - T_t)i_t$), as Oikarinen (2007, p. 107) recommends.

Unfortunately, there is not data in Finland to cope with ${}^{\lambda_t}/_{\mu_c}$, Oikarinen (2009) believes that its effects are found within the household debt data in our model. In this thesis this component is considered separately from U_t. More detailed information about it is found in section "loan-to-GDP ratio" (3.3.6).

DiPasquale and Wheaton (1994) and Oikarinen (2009) recommend the adoption of adaptative expectations perspective (backward looking) in order to estimate the expected real appreciation component $E(P_{t+1} - P_t)$, known as well as "capital gains". Oikarinen draws attention either on its computational problems and on the necessity of using real averages instead of nominal ones since the inflation rates in the present times[68] are lower than in the last decades of 1900s (p. 129).

The risk premium (γ) is set as a time-invariant 2%, in the same direction than Oikarinen (2009) and Himmelberg et al. (2005)

Unfortunately, $E(P_{t+1} - P_t)$ and γ are finally removed from the equation (55). Estimation without them generates better results. Hence, they are not considered in the final analysis part.

[68] Expected real appreciation follows Himmelberg et al. (2005).

3.3.6 Loan-to-GDP ratio (L_t)

Following Hort (1998) and Oikarinen (2009 & 2012), we included a variable to measure credit availability. It is not included in the "original" life-cycle model but according to Oikarinen (2009) is a good proxy for λ_t/μ_c. Moreover, it correlates with credit constraints.

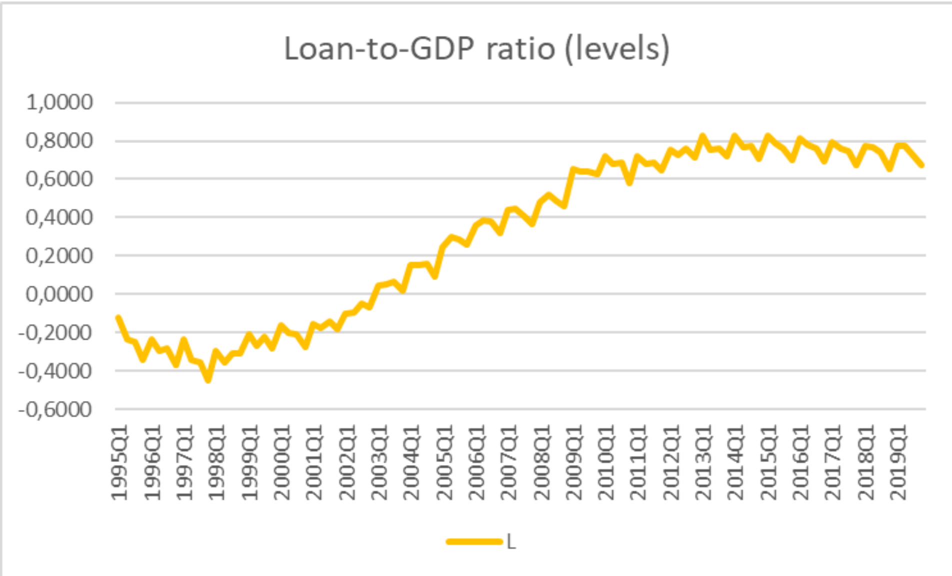

Figure 12. Evolution of loan-to-GDP ratio (levels) in Finland 1995Q1-2019Q4. (Sources: Tilastokeskus and Suomen Pankki)

The data on month-end outstanding mortgage loan stock granted by Finnish financial institutions to residents is published by the Suomen Pankki[69]. The unit is in million current euros. The outstanding mortgage loan stock has been transformed into quarterly data and deflated by CPI (2010=100). Hort (1998) uses the net lending per household and Oikarinen (2012) makes use of the outstanding mortgage loan stock of households.

[69] From https://www.suomenpankki.fi/en/Statistics/mfi-balance-sheet/tables/rati-taulukot-en/lainat_kotitalouksille_en/

As numerator, we use the net disposable income of Finland in million euros (after taxes and transfers). The data is quarterly released by Tilastokeskus and sets 2010 as reference year.

Unfortunately, we did not find enough disaggregated regional-level data covering the period 1995-2019. Data illustrates the situation at national level, it is however reasonably to assume that CRH outstanding mortgage loan stock development resembles the one described by national data. According to Tilastokeskus, the household-dwelling units held liabilities, generally speaking, during 2019 about 220 per cent of their disposable income (Statistics Finland, indebtedness). We noticed the same trend in our adjustments.

4 DATA ANALYSIS

This chapter presents the empirical findings from the Johansen cointegration analysis. It begins with the unit root test, the optimal lag length determination. Once it is established, we proceed with the cointegration test, we continue by estimating the corresponding VECM and determining the long-term and short-term relationships among house prices and their fundamentals. Finally, the required residual analyses are performed, we test for residual autocorrelation, residual normality and autoregressive conditional heteroskedasticity.

The results are consistent with economic theory. Only three models are reported in this thesis, the "best models". They are estimated as a loglinear approximation. The first model, the parsimonious model, includes only Y_t as independent variable. The second model has Y_t and L_t as independent variables. Finally, the third model incorporates Y_t, L_t and U_t. Econometric analysis has been performed with EViews 11 Student Version software package.

Our analyses include 3 quarterly centred seasonal dummies: S_2, S_3 and S_4 (to cater or seasonal effects on quarter 2, 3 and 4).

4.1 Unit root test

A necessary precondition for a stationary long-term relationship to exist, and therefore for cointegration, is that time series at levels exhibit unit root[70] (I(1)), and thus their linear combination is stationary.

In the first place the stationarity of time series is examined visually (see i.e., Figures 7-12). Time series are non-stationary at levels. The visual inspection is followed by the ADF test. Table 1 summarises the unit root statistics. P_t at levels is modelled with constant and without it. Y_t, U_t and L_t are modelled with a constant in levels and H_t and CC_t without constant. Since variables are in natural logarithms, but U_t, following

[70]At least, I(1) or in some special cases I(2).

Oikarinen (2007) assumption no further check was incorporated in this analysis for trend. The criterion employed is the SBC.

All times series but P_t exhibits certainly non-stationarity at levels, clearly I(1). For P_t, when including a constant, it is non-stationary at 10% significance level. Although when modelled without constant it is clearly non-stationary (see Table 1). The order of integration of U_t is ambiguous[71]: ADF test suggests that user cost is non-stationary at levels, but it reveals close to stationarity.

As expected, every time series exhibits stationarity at differences. P_t, Y_t, L_t, H_t and U_t are modelled with a constant. The ADF test for CC_t in differences reveals non significance of the constant at 10%. Hence, all variables can be treated as non-stationary at levels[72] and stationary in differences. Cointegration analysis is viable. For visual inspection of the time series, the figures can be seen in Appendix 2 (Figures 13-18).

4.2 Models

We have tested different combinations and possibilities and only three models have produced satisfactory results. Based on equation (59) the models estimated are:

$$P_t = \beta_1 Y_t + \varepsilon_t \tag{63}$$

$$P_t = \beta_1 Y_t + \beta_2 L_t + \varepsilon_t \tag{64}$$

$$P_t = \beta_1 Y_t + \beta_2 L_t + \beta_3 U_t + \varepsilon_t \tag{65}$$

Equation (63) is the Model 1; equation (64) is the Model 2 and (65) represents the Model 3. All the models are estimated with some unrestricted constant within the cointegration equation and the VAR equation and incorporate centred seasonal dummies.

[71] Oikarinen (2009) finds similar issues for U_t.

[72]Although U_t rises justifiable doubts.

Table 1. Augmented Dickey-Fuller test results

Variable	Level (lags)	1 %	5 %	10 %	Differences (lags)	1 %	5 %	10 %
Real Housing prices (P)	-2.942** (1)c	-3.498	-2.891	-2.583	-5.450*** (0)c	-3.498	-2.891	-2.583
Real Housing prices (P)	2.090 (1)	-2.589	-1.944	-1.615				
Real aggregate income (Y)	-2.250 (0)c	-3.498	-2.891	-2.583	-11.497***(0)c	-3.498	-2.891	-2.583
Real user costs (U)	-2.223(0)c	-3.498	-2.891	-2.583	-4.651***(4)c	-3.501	-2.893	-2.583
Loan-to-GDP ratio (L)	-1.695(4)c	-3.501	-2.892	-2.583	-3.283***(3)c	-3.501	-2.892	-2.583
Housing stock (H)	0.807(4)	-2.590	-1.944	-1.615	-4.364(3)c	-3.501	-2.892	-2.583
Construction Costs (C)	1.831 (0)	-2.589	-1.944	-1.615	-7.639***(0)	-2.589	-1.944	-1.615

*, ** and *** denotes for statistical significance at 10%, 5% and 1% respectively. c denotes for a constant. Based on SBC criterion. Max lag=4.

In the same vein than Oikarinen (2009), time series H_t and CC_t are not included in any analysis, and therefore no supply-side determinant seems relevant. The results obtained in the cointegration analysis are not satisfactory: either cointegration is not found or they are statistically insignificant at any significance level and their inclusion does not improve the explanatory power of the models. We suspect that H_t introduces multicollinearity in the models reducing the precision of the estimated coefficients, and therefore p-values are not reliable when identifying which variables are statistically significant. Observing Figure 9, we are of the same opinion than Oikarinen (2009), the construction costs show sufficient flat trend and therefore their effect on House prices can be ignored. Hence, $\beta_4 = \beta_5 = 0$.

4.3 Optimal lag length

Once the non-stationarity is evaluated, the next step is to set the optimal lag length for each of the three feasible models. Using the whole sample, a VAR in levels is estimated for each model and the lag length is chosen according to the criteria. For all models, HQ and SBC criteria are compared and preferred to AIC to determine correct lag length. For Model 1 SBC and HQ determine two lags. For Model 2, SBC determines only one lag.

According to SBC criterion, Model 3 only needs one lag. Following Oikarinen (2007), an additional lag length is included, because the Lagrange Multiplier (LM) test rejects the hypothesis of no serial autocorrelation (HQ criterion suggests 2 lags). With two lag length, Model 3 passes the LM test for autocorrelation. More detailed information about each of the two criteria can be found in the tables from Appendix 3 (Tables A3.1-A3.3).

4.4 Cointegration test

Once the lag length is determined by information criteria, we can proceed with the Johansen cointegration test. Johansen Cointegration test is a procedure for testing cointegration rank and to determine the correct parametrization of the VECM. Hence, using the corresponding lag length we include an intercept within de cointegration equation, allowing for linear deterministic trend in the data.

Johansen Cointegration test will determine the cointegrating rank of the VECM and therefore the existence of matrices α and β[73].

For all three models both trace test and eigenvalue test support the existence of 1 cointegration equation. In all three models we are rejecting the null hypothesis of no cointegration at 5% significance level, and therefore at least 1 cointegration equation exists. In Table 2 all the results are summarized, the trace statistics, eigenvalue statistics and the corresponding p-values for each one of them.

Table 2. Johansen trace test and eigenvalue test

	Model 1			
Hypothesis	r=0	r≤1	r≤2	r≤3
Trace statistics	17.073*	0.518	-	-
p-value**	0.029	0.471	-	-
Max-Eigen statistics	16.555*	0.518	-	-
p-value**	0.0213	0.472	-	-
	Model 2			
Hypothesis	r=0	r≤1	r≤2	r≤3
Trace statistics	38.005*	8.802	1.398	-
p-value**	0.005	0.384	0.237	-
Max-Eigen statistics	29.202*	7.404	1.398	-
p-value**	0.003	0.443	0.237	-
	Model 3			
Hypothesis	r=0	r≤1	r≤2	r≤3
Trace statistics	76.619*	25.842	11.516	1.620
p-value**	0.000	0.134	0.182	0.203
Max-Eigen statistics	50.777*	14.326	9.896	1.620
p-value**	0.000	0.339	0.219	0.203

* denotes rejection of hypothesis at 5% significance level. ** MacKinnon-Haug-Michelis (1999) p-values. Exogenous series: S2 S3 S4.

Given that Johansen Cointegration test (both trace and eigenvalue statistics) supports, three different VECM are estimated to perform empirical analysis.

[73] In the parsimonious case, the vectors α and β.

4.5 Model 1: ln P = ln Y

4.5.1 Long-term relationship

Results suggest the existence of one cointegration equation with 2 lags. Normalizing the cointegrating equation with respect to P_t, the long-term equilibrium relationship is:

$$P_t = -0.354 + \underset{(0.139)}{0.688} Y_t \tag{66}$$

This model represents the parsimonious model. In (66) there is a positive long-term relationship between the real aggregate income and the real house prices. Model includes an unrestricted constant. The standard error is in parentheses.

The sign of the estimated parameter is according to the theory, and the value of the coefficient is 0.688. It is in line with Bourassa et al. (2019) findings. They estimate a parsimonious model for Helsinki, although their income elasticity is just slightly higher, 0.737.

According to the adjusted coefficient of determination (Adj. R^2) the explanatory power of the estimated model explains about 44.6% of the house price changes. The income elasticity concludes that 1% increase in aggregate income causes real prices to change around 0.69%. A detailed description of Model 1 can be found in the Appendix 4 Table A4.1.

4.5.2 Short-term relationship

We can rewrite (66) as the error correction equation, as in (61):

$$\epsilon_{t-1} = 1.0000 P_{t-1} - \underset{(0.139)}{0.688} Y_{t-1} + 0.354 \tag{67}$$

where ϵ_{t-1} the one term lagged deviation.

Table 3. Error Correction model

(Model 1)	D(P)	Stand. Error
α	-0.054***	(0.014)
P(-1)	0.517***	(0.106)
P(-2)	-0.010	(0.100)
Y(-1)	-0.047	(0.096)
Y(-2)	-0.074	(0.095)
Constant	0.016***	(0.003)
S2	-0.007	(0.005)
S3	-0.014***	(0.005)
S4	-0.018***	(0.004)
Adjusted R^2	0.446	
J-B	0.000	
LM(1)	0.311	
LM(2)	0.264	
LM(heter)	0.001	
Akaike criterion	-10.730	
Schwarz criterion	-10.199	

*, ** and *** denotes for statistical significance at 10%, 5% and 1% respectively. LM(1) and LM(2) are the Lagrange-multiplier test for residual autocorrelation at lag 1 and lag 2. J-B represents the Jarque-Bera test for residual normality. LM(heter) is the joint Lagrange-multiplier for heteroskedasticity in residuals.

Table 3 summarises our findings. The speed of adjustment, has the expected negative sign, supporting the hypothesis of long-term adjustment towards the equilibrium. The speed of adjustment of house prices indicates convergence towards the equilibrium about 5.4% per quarter (about 23% a year). These results are in the same line than Barot and Takala (1998) who observe price adjustment about 4.4% a quarter in Finland. Results seem to confirm that prices levels in the CRH adjust sluggishly towards the long-term equilibrium level.

Granger causality test supports the hypothesis of short-term causal effects of the lagged house prices on present house prices. Unfortunately, test do not support real aggregate income having a short-term causal effect on house prices. (Table A4.7) The joint test from the Wald test does not support aggregate income as house price determinant in the short run. The joint Wald test supports the inclusion of the centred seasonal dummies in the model (Appendix 4-Table A4.2, Table A4.3.1 and Table A4.3.2).

4.6 Model 2: ln P= ln Y + ln L

4.6.1 Long-term relationship

The long-term equilibrium relationship is:

$$P_t = -0.493 + 0.697Y_t \; + 0.237L_t \qquad (68)$$
$$(0.156) \quad (0.081)$$

In the long-run both real aggregate income and loan-to-GDP ratio are statistically significant at 1%. As in Oikarinen (2009) aggregate income is restricted as weakly exogenous. The standard errors are in parentheses.

The income elasticity concludes that 1% increase in aggregate income causes real prices to change around 0.70%. The sign of the estimated parameter is according to the theory, positive, and the value of the coefficient is 0.697. This is in line with the multivariate model in Bourassa et al. (2019) for Helsinki, 0.910; or with the model 6 in Oikarinen (2005), where he estimates an elasticity of 0.829. Model 2 obtains similar results as Kuismanen et al. (1999) as well, with income elasticity 0.808, although this coefficient represents the lagged value of income.

Compared with other relevant international findings, Meese and Wallace (2003) obtain a similar coefficient for income elasticity of 0.646. Hort (1998) for Sweden estimates an income elasticity of 0.965. Quite similar results are obtained by Abraham and Hendershott (1996) for U.S. as well (0.774).

Loan-to-GDP elasticity supports that 1% increase in L_t causes real prices to change around 0.24%. The value of the coefficient of L_t resembles the one estimated by Oikarinen (2009), 0.355; and specially the one estimated in Oikarinen (2005, model 6), 0.282; and the one obtained in Oikarinen (2012), 0.271. The inclusion of L_t unveils information about the borrowing capacity of households, as suggests Oikarinen (2009).

According to the adjusted coefficient of determination (Adj. R^2) the explanatory power of the estimated model explains about 40.8% of the house price changes. A detailed description of Model 2 can be found in the Appendix 5-Table A5.1.

4.6.2 Short-term relationship

In this section the short-term relationship among house prices and their fundamentals and the speed of adjustment of the parameter are analysed. Rewriting (68) as the error correction equation, as in (61):

$$\epsilon_{t-1} = 1.0000P_{t-1} - \underset{(0.156)}{0.697}Y_{t-1} - \underset{(0.081)}{0.237}L_{t-1} + 0.493 \qquad (69)$$

Table 4. Error Correction model

(Model 2)	D(P)	Stand. Error
α	-0.045**	(0.023)
P (-1)	0.617***	(0.088)
Y(-1)	-0.033	(0.097)
L(-1)	0.006	(0.057)
Constant	0.016***	(0.005)
S2	-0.012	(0.011)
S3	-0.018***	(0.005)
S4	-0.019***	(0.006)
Adjusted R^2	0.408	
J-B	0.000	
LM(1)	0.532	
LM(heter)	0.000	
Akaike criterion	-15.045	
Schwarz criterion	-14.333	

*, ** and *** denotes for statistical significance at 10%, 5% and 1% respectively. LM(1) is the Lagrange-multiplier test for residual autocorrelation at lag 1. J-B represents the Jarque-Bera test for residual normality. LM(heter) is the joint Lagrange-multiplier for heteroskedasticity in residuals.

Empirical findings (Table 4) support the hypothesis that there is adjustment towards the long-term equilibrium. The α coefficient is statistically significant at 5% significance level. The speed of adjustment of house prices indicates convergence towards the equilibrium about 4.5% per quarter (Table A5.1). These results are in the same line than Barot and Takala (1998) (4.4% a quarter).

As estimated in Model 1 house prices show sluggishness towards their adjustment towards the long-term equilibrium level, they converge slowly. Lagged values of house prices are the short-term causal effects of house prices changes according to Granger-causality test from Table A5.6 (for more detail about the short-term part of the model see e.g., Appendix 5, Table A5.1 and Table A5.2.1). The inclusion of seasonal dummies is relevant in this model (Table A5.2.2).

4.7 Model 3 ln P = ln Y + ln L + ln U

4.7.1 Long-term relationship

The long-term equilibrium relationship is:

$$P_t = 0.905 + \underset{(0.113)}{0.492}Y_t + \underset{(0.057)}{0.302}L_t \tag{70}$$

In the long-run both real only aggregate income and loan-to-GDP ratio are statistically significant at 1%. Real cost of ownership is non-significant and restricted from the long-term model (as in Oikarinen 2009, $\beta_3 = 0$). The standard errors are represented in parentheses.

The sign of the estimated parameters is according to the theory, positive for aggregate income and for loan-to-GDP ratio. The value of the coefficient of income is 0.492. The income elasticity concludes that 1% increase in aggregate income causes real prices to change around 0.49%. The results for Y_t and L_t are in line with Oikarinen (2009), with an estimated income elasticity of 0.356, loan-to-GDP ratio elasticity of 0.355 and intercept 1.22. Although this analysis does not support the exogeneity of real income (p-value too close to 5% level).

The estimated coefficient of the income elasticity is in the same line as other relevant international and domestic research such as e.g., Pere and Takala (1991), 0.472 or Capozza et al. (2004), 0.45.

Loan-to-GDP elasticity supports that 1% increase in L_t causes real prices to change around 0.30%. The value of the coefficient of L_t resembles the one previously estimated in Model 2. The value of the coefficient is in the same line than other works from Oikarinen[74] with elasticities ranging from 0.271 to 0.356. As already mentioned in Model 2, L_t reveals information about the borrowing capacity of households in the CRH. (Oikarinen 2009)

The no inclusion of U_t in the long-term model is not a big surprise since ADF test revealed ambiguous results rising doubts about its stationarity. Hence, real cost of ownership is restricted from the long-term model. Oikarinen (2009) excludes U_t from the long-term model as well arguing that it is "(at least close to) stationarity" (p. 133).

According to the adjusted coefficient of determination (Adj. R^2) the explanatory power of the estimated model is about 51.4% of the house price changes. A more detailed description of Model 3 can be found in the Appendix 6-Table A6.1.

4.7.2 Short-term relationship

Rewriting (70) as the error correction equation:

$$\epsilon_{t-1} = 1.0000P_{t-1} - \underset{(0.113)}{0.492}Y_{t-1} - \underset{(0.057)}{0.302}L_{t-1} \tag{71}$$

Empirical findings (Table 5) support the hypothesis that there is adjustment towards the long-term equilibrium. The speed of adjustment of house prices indicates convergence towards the equilibrium about 6.8% per quarter and about 30% a year (Table A6.1). House prices are sluggish, detecting thus inefficiency of housing markets in CRH. These results are in the same line than Oikarinen (2009), who estimates a convergence about 6.4% a quarter. These results are in the same line than Oikarinen (2007), 7.2% per quarter or Takala and Pere (1991), 6.9%-7.4% each quarter.

[74] 2005, model 6, 2007, 2009 and 2012.

Table 5. Error Correction model

Model 3	D(P)	Stand. Error
α	-0.068***	(0.025)
P(-1)	0.440***	(0.116)
P(-2)	0.050	(0.102)
Y(-1)	-0.008	(0.095)
Y(-2)	-0.075	(0.090)
L(-1)	0.001	(0.055)
L(-2)	0.135**	(0.057)
U(-1)	-1.513***	(0.636)
U(-2)	-0.909	(0.648)
Constant	0.017***	(0.004)
S2	-0.005	(0.011)
S3	-0.033***	(0.008)
S4	-0.015***	(0.005)
Adjusted R^2	0.514	
J-B	0.000	
LM(1)	0.210	
LM(2)	0.093	
LM(heter)	0.005	
Akaike criterion	-24.352	
Schwarz criterion	-22.866	

*, ** and *** denotes for statistical significance at 10%, 5% and 1% respectively. LM(1) and LM(2) are the Lagrange multiplier test for residual autocorrelation at lag 1. J-B represents the Jarque-Bera test for residual normality. LM(heter) is the joint Lagrange-multiplier for heteroskedasticity in residuals.

Lagged values of house prices (with positive sign) and of U_t are the short-term Granger-causal effects of house prices changes in this model (Table A6.6). The sign of U_t is, as expected, negative and therefore according to what theory predicts. Joint Wald test supports the short-term causality of loan-to-GDP ratio, although only second-lag is statistically significant (Table A6.2.1), we suspect that L_t might have mid and long-term effects on house prices, but not sort-term ones. In section 5, these findings are discussed (for more detail about the short-term part of the model see e.g., Appendix 6, Table A6.1 and Table A6.2).

Based on Akaike and SBC information criteria, Model 3 has the lowest values. Based on Adj. R^2, Model 3 performs better, with higher explanatory power. Therefore, Model 3 is set a preferred. (Table 3, Table 4, and Table 5)

4.8 Diagnostics checks

Appendices 4, 5 and 6 include tables with results from LM test for serial autocorrelation, the residual test for normality and the test for heteroskedasticity. Main results can be found in Table 3, Table 4, and Table 5 as well.

According to LM(1) and LM(2) test, Model 1 does not show residual autocorrelation at 5% significance level, we cannot reject the null hypothesis of no serial correlation (p-values for lag 1, 0.107 and for lag 2, 0.732) Neither LM (1) test for Model 2 shows residual autocorrelation at 5% significance level (p-value 0.073).

Model 3, according to LM(1) and LM(2) test, does not show residual autocorrelation at 5% significance level, we cannot reject the null hypothesis of no serial correlation (p-values for lag 1, 0.210 and for lag 2, 0.093) Hence, the three VECM provide an accurate representation of the DGP (Tables A4.4, A5.3, and A6.3).

For Model 1, Model 2, and Model 3 Jarque-Bera test, as expected, residuals are non-normally distributed (Tables A4.5, A5.4, and A6.4) and show some heteroskedasticity (Tables A4.6, A5.5, and A6.5).

No further adjustment is made since the robustness of the cointegration rank tests is not considerably affected by medium ARCH effects. (Oikarinen 2007, Cavaliere, Rahbek and Taylor 2010) Authors such as Oikarinen (2009) allow for heteroskedasticity in their econometric analyses.

5 SUMMARY

Model 3 is the best performing and therefore describes the long-term and short-term dynamics of the housing market in the CRH. Our model resembles the model estimated in Oikarinen (2009), although income elasticity coefficient slightly diverges from the one estimated by Oikarinen. This coefficient caters for both, the demographic factors and permanent income. Two suitable explanations for this divergence: on the one hand, internal and external migration movements have concentrated population in urban areas (especially in CRH) and on the other hand, both studies face different sample periods: Oikarinen (2009) covers a period with substantial instability (i.e., financial deregulation during the 1980s and the economic crisis of the 1990s). Finland, and specially CRH, have enjoyed of sustained growth that explains the upward trend of aggregate income for most of the sample period. Hence, as our research suggests both determinants have a positive relationship with house prices.

Our “best model” suggests the existence of a long-term relationship among house price disposable permanent income, demographic factors, and loan-to-GDP. Whereas in the short run, lagged prices and lagged real user costs show short-term causality on house prices. These results are consistent with previous results from Oikarinen (2009).

Results from previous research, both domestic and international, support our results. Our estimated income elasticities (0.492-0.697) are in the line of authors such as Abraham and Hendershott (1996), 0.774; Kuismanen et al. (1999), 0.808; Meen and Wallace (2003), 0.646; Oikarinen (2009), 0.356; and Bourassa et al. (2019), 0.737. Our estimated loan-to-GDP ratio elasticity (0.237-0.302) is, as well, in line with previous research conducted in Finland (see e. g. Oikarinen 2005, Oikarinen 2009 and Oikarinen 2012), and is quite robust.

Although loan-to-GDP ratio it is not included in the “original” life-cycle model, according to Oikarinen (2009) it captures the credit availability/liquidity constraints. During the sample period, the outstanding mortgage loan stock has constantly increased, and therefore no substantial credit constraints are observed. An explanation is given by Oikarinen (2007): when households experience periods with easy access to borrowing along with low levels of income uncertainty, they change their perceptions on permanent income, smoothening their long-term consumption path and therefore their liquidity constrains change. This translates into more willingness to get into debt and therefore increasing their demand for housing. According to our findings, loan-to-GDP ratio does not affect house prices in the short run, we believe that the lack of credit constraints has no short-term effect on house prices. We believe its effect are mid-term and long-term, although we should conduct more research (for instance through impulse response function and variance decomposition analyses).

The real user cost, in the form of the real-after tax mortgage rate is excluded from the long-term equation. Oikarinen (2009) offers us a feasible answer, “user cost is mean-reverting” (p. 133): he states that “housing demand of forward-looking agents with long planned holding periods of housing should not react strongly to changes in it” (p. 133). The whole sample period has enjoyed from decreasing inflation rates combined with decreasing average mortgage rates. The overall effect has driven real user costs down, making more attractive the access to homeownership (although during 2000-2001 and 2006-2008 the real user costs increased). Model 3 finds that due to its mean-reverting nature, its downward trend, does not translate into long-term effects, but in the short run households are taking advantage of the reduced cost of borrowing, increasing their demand for housing, and therefore increasing house prices in the short run. (the negative relationship ensures that decreasing real user costs result in increasing real house prices). (Oikarinen 2007) We suspect that rational agents are myopic (irrational) in the short run, driven by individual “speculative behaviour”, and the seed for speculative bubbles. According to Himmelberg (2005), agents have a distorted perception on user costs and end up paying “too much”. (See as well e.g., DiPasquale & Wheaton 1994 and Abraham & Hendershott 1996). Oikarinen (2009) suggests that information inefficiencies could be behind of such behaviours.

According to our analysis, the house prices are far from following a random walk process. As Poterba et al. (1991) state, housing market is predictable based on the positive serial price correlation. The three models show short-term causality between lagged values of house prices and current prices. Other authors prove our thesis (see, e.g., Case & Schiller 1988 (only when there is a positive relationship), Capozza et al. 2004 or Steiner 2010)

Another important finding is the “inefficiency” of the Helsinki housing market. Perfect market clearing conditions have to be relaxed, according to DiPasquale and Wheaton (1994) “may not be rational” (p. 6). Reasons describing why market prices are much sluggish than expected and why real estate markets respond slower than theory predicts are due to market-related characteristics, to the product heterogeneity, to time-consuming search or mismatches between new construction and demand. (DiPasquale & Wheaton 1994) This is confirmed as well by relevant previous research, e.g., Mankiw and Weil (1989), Weil and Shiller (1991), Case and Schiller (1990), Drake (1993), DiPasquale and Wheaton (1994), Oikarinen (2009), Oikarinen (2012) or Oikarinen (2014).

The results from Model 3 suggest that house prices converge sluggishly towards the long-term equilibrium level, about 6.8% ($\alpha = -0.068$). Other estimated models show a speed of quarterly adjustment within 4.5% and 5.4%. Results are in the line of previous research conducted in Finland (see e.g., Takala & Pere 1991, Barot & Takala 1998 or Oikarinen 2009).

Unfortunately, supply-side fundamentals in CRH do not reflect the existence of a long-term relationship with real house prices, in the line with other research (see e.g., Oikarinen 2009 or Bourassa et al. 2019). The total of building completions, used as a proxy to cater with the evolution of housing stock (H_t) has proven problematic. This data, almost certainly, induces multicollinearity in our analyses: for instance, the sign of some variables is reversed. Some p-values have shown not reliable eroding the precision of some coefficients. A better proxy would be some related to the number of dwellings, but time series from Tilastokeskus was not long enough to complete the whole sample period.

Real house prices show a rising trend. However, this increase seems not to be linked with construction costs (CC_t) since they stay constant during the whole sample period. Theoretical framework from DiPasquale and Wheaton (1992) suggests that market prices and construction costs, may diverge in the short-term but converge towards each other in the long-term. Consequently, any gap between them generates construction activity. We partially agree: the case of CRH confirms that the existence of such a gap boosts construction activity, but data from the last 30 years does not support such convergence between them. House prices have increased at a faster pace than construction costs during the whole sample period.

Although, we must not be surprised of their irrelevance, on addition to Oikarinen (2009) or Bourassa et al. (2019) findings, Meen (2002) already highlights their irrelevance as regressors in most research in U.K. Even Riddle (2004) finds no long-term causality between house prices and construction costs and neglects their short-term causality in U.S..

Our analysis could be completed by incorporating and conducting Impulse Response Function and Variance Decomposition analyses: studying how the fundamentals and house prices react to different shocks and decompose each shock contribution over the other variables over time. Since most of the literature reviewed focus attention on the CRH and other big urban areas, such as Turku, Tampere or Oulu, further research is encouraged on long-term and short-turn determinants of house prices in other regions from Finland that have captured less attention, such as countryside or medium cities. Further research could be extended, both at micro-level and/or at macro-level, covering cities such as Vantaa or Espoo/Kauniainen and/or extending the analysis to the greater Helsinki area (Tuusula, Järvenpää, Mäntsälä, Kerava or Riihimäki). Professor Oikarinen in a private conversation stressed on the complexities of studying secondary cities within a metropolitan area, such as Vantaa, since most of the analyses are "conducted at the metropolitan area level or concerning the central city of a metro area"[75]. He suggested these issues could be avoided by using aggregated data on CRH on population and income. Hence, the whole regional housing demand in the region, relevant to price development, is captured, although some of the specific regional developments are lost when using such homogenised data.

[75] According to his own words.

REFERENCES

Abraham, J. & Hendershott, P. H. (1996). “Bubbles in Metropolitan Housing Markets.” Journal of Housing Research 7, 191–207.

Act Amending the Income Tax Act, 8.12.1995/1389. Retrieved from https://www.edilex.fi/smur/19921535

Act Amending the Income Tax Act, 23.12.1999/1225. Retrieved from https://www.edilex.fi/smur/19921535

Act Amending the Income Tax Act, 30.07.2004/716. Retrieved from https://www.edilex.fi/smur/19921535

Act Amending the Income Tax Act, 29.12.2011/1515. Retrieved from https://www.edilex.fi/smur/19921535

André, C. and García, C. (2012). "Housing Price and Investment Dynamics in Finland", OECD Economics Department Working Papers, No. 962, OECD Publishing, Paris, https://doi.org/10.1787/5k98rwldjr44-en.

Barot, B. & Takala, K. (1998). “House prices and inflation: a cointegration analysis for Finland and Sweden.” Research Discussion Papers 12/1998, Bank of Finland

Bourassa, S.C., Hoesli, M. & Oikarinen, E. (2019). Measuring House Price Bubbles. Real Estate Economics, 47: 534-563. https://doi.org/10.1111/1540-6229.12154

Bulan, L., Mayer C. & Somerville, C. T. (2009). “Irreversible investment, real options, and competition: Evidence from real estate development.” Journal of Urban Economics 65 (3): 237–51.

Capozza, D. R. & Li, Y. (1994). The Intensity and Timing of Investment: The Case of Land. The American Economic Review, 84(4), 889-904. Retrieved March 11, 2021, from http://www.jstor.org/stable/2118036

Capozza, D. R., Hendershott, P. H., Mack, C., and Mayer, C. J. (2002). “Determinants of Real House Price Dynamics,” NBER Working Paper 9262, Cambridge, Massachusetts.

Capozza, D. R., Hendershott, P. H., & Mack, C. (2004). An anatomy of price dynamics in illiquid markets: Analysis and evidence from local housing markets. Real Estate Economics, 32(1), 1-32. https://search.proquest.com/scholarlyjournals/anatomy-price-dynamicsilliquidmarketsanalysis/docview/211152401/se2?accountid=13031

Case, K. E., & Shiller, R. J. (1989). The Efficiency of the Market for Single-Family Homes. The American Economic Review, 79(1), 125-137. Retrieved January 21, 2021, from http://www.jstor.org/stable/1804778

Case, K. E., & Shiller, R. J. (1990). Forecasting prices and excess returns in the housing market. AREUEA Journal, 18(3), 253. Retrieved from https://search.proquest.com/scholarly-journals/forecasting-prices-excess-returns-housing-market/docview/211136431/se-2?accountid=13031

Cavaliere, G., Rahbek, A., & Taylor, A. M. R. (2010). Cointegration rank testing under conditional heteroskedasticity. Econometric Theory, 26(6), 1719-1760. doi:http://dx.doi.org/10.1017/S0266466609990776

Chen, M. & Patel, K. (1998). "House Price Dynamics and Granger Causality: An Analysis of Taipei New Dwelling Market," International Real Estate Review, Global Social Science Institute, vol. 1(1), pages 101-126

Chow, G. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica, 28(3), 591-605. doi:10.2307/1910133

Colwell, P. F. (2002). Tweaking the DiPasquale-Wheaton Model, Journal of Housing Economics, 11 (1), p. 24-39. https://EconPapers.repec.org/RePEc:eee:jhouse:v:11:y:2002:i:1:p:24-39.

DeSalvo, J. S. (2017). Teaching the DiPasquale-Wheaton model. Journal of Real Estate Practice and Education, 20(1), 1-25. Retrieved from https://search.proquest.com/tradejournals/teachingdipasqualewheatonmodel/docview/1933857042/se-2?accountid=13031

DiPasquale, D. & Wheaton, W.C. (1992). The Markets for Real Estate Assets and Space: A Conceptual Framework. Real Estate Economics, 20: 181-198. https://doi.org/10.1111/1540-6229.00579

DiPasquale, D. & Wheaton, W. C. (1994). Housing Market Dynamics and the Future of Housing Prices, Journal of Urban Economics, 35, issue 1, p. 1-27

Drake, L. (1993). Modelling UK house prices using cointegration: an application of the Johansen technique. Applied Economics, 25(9), 1225. https://doi-org.pc124152.oulu.fi:9443/10.1080/00036849300000183

Dye, R. & Mcmillen, D. (2007). Teardowns and Land Values in the Chicago Metropolitan Area. Journal of Urban Economics. 61. 45-63.

Enders, W. (2014). Applied Econometric Time Series."4th ed". New York (US): University of Alabama.

Engle, R., & Granger, C. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica, 55*(2), 251-276. doi:10.2307/1913236

Fisher, J.D. (1992). Integrating Research on Markets for Space and Capital. Real Estate Economics, 20: 161-180. https://doi.org/10.1111/1540-6229.00578

Gimeno, R. & Martínez-Carrascal, C. (2010). The relationship between house prices and house purchase loans: The Spanish case. Journal of Banking & Finance 34: 1849-1845.

Giussani, B. & Hadjimatheou, G. (1990). House prices: an econometric model for the UK, *The APEX Centre Economics Discussion Paper* 90/1, Kingston Polytechnic, London.

Hamilton, B. W. & Schwab, R. M. (1985). Expected appreciation in urban housing markets, Journal of Urban Economics, 18, issue 1, p. 103-118, https://EconPapers.repec.org/RePEc:eee:juecon:v:18:y:1985:i:1:p:103-118.

Hansen, P. (2003). “Structural changes in the cointegrated vector autoregressive model” Journal of Econometrics 114(2) 261–295

Helin, H. (2007). Kiinteistövero: Suuret kaupungit 1993–2007. Helsingin kaupungin tietokeskus. Tutkimuskatsauksia 2007, 5. ISSN 1796-7236. ISBN 978-952-473-970-2. Retrieved from: https://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/07_11_01_Tutkkats_5__Helin.pdf

Higgins, R. S. (1972). “Microeconomic Foundations of the Flexible Accelerator,” Western Economic Journal 10(1), 19–32.

Himmelberg, C., Mayer, C. & Sinai, T. (2005). Assessing high house prices: bubbles, fundamentals and misperceptions. Journal of Economic Perspectives 19, 67–92.

Hort, K. (1998) The Determinants of Urban House Price Fluctuations in Sweden 1968–1994, Journal of Housing Economics, 7, issue 2, pp. 93-120

Hott, C. (2009). “Explaining house price fluctuations,” Swiss National Bank Working Papers, (5).

Johansen, S. (1988). Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control, 12, issue 2-3, p. 231-254

Johansen, S., & Juselius, K. (1990). “Maximum likelihood estimation and inference on cointegration—with applications to the demand for money,” Oxford Bulletin of Economics and Statistics 52, pp. 169–210.

Kaleva, H., Niemi, J. & Törnroos, J. (n.d.) Kiintestön omistamisen verotus. Raportti. KTI Kiinteistötieto Oy. Retrieved from: https://kti.fi/wp-content/uploads/Kiinteist%C3%B6n-omistamisen-verotus.pdf

Kearl, J. R. (1979). Inflation, Mortgages, and Housing, Journal of Political Economy, 87, issue 5, p. 1115-38.

Kuismanen, M., Laakso, S., & Loikkanen, H. A. (1999). Demographic factors and the demand for housing in the Helsinki Metropolitan Area. Valtion taloudellinen tutkimuskeskus VATT.

Loikkanen, H. A., & Lönnqvist, H. (2007). Metropolitan housing markets—The case of Helsinki. In European metropolitan housing markets (pp. 63-84). Springer, Berlin, Heidelberg.

Linneman, P. (1986). An Empirical Test of the Efficiency of the Housing Market. Journal of Urban Economics 20:140–54.

Lütkepohl, H. & Krätzig, M. (2004). Applied Time Series Econometrics, Cambridge University Press, https://EconPapers.repec.org/RePEc:cup:cbooks:9780521839198

Lütkepohl, H. & Krätzig, M. (2005). VECM Analysis in JMulTi July 5 , 2005. Retrieved from https://www.semanticscholar.org/paper/VECM-Analysis-in-JMulTi-July-5-%2C-2005-L%C3%BCtkepohlKr%C3%A4tzig/f8163639f40df78314955a3c0f62d6c5ca1b3bf3

Malpezzi, S. (1999). A Simple Error Correction Model of House Prices, Journal of Housing Economics, 8, issue 1, p. 27-62, https://EconPapers.repec.org/RePEc:eee:jhouse:v:8:y:1999:i:1:p:27-62.

Mankiw, N. G. & Weil, D. (1989), The baby boom, the baby bust, and the housing market, Regional Science and Urban Economics, 19, issue 2, p. 235-258

McCarthy, J. & Peach, R. W. (2002). Monetary policy transmission to residential investment. Economic Policy Review - Federal Reserve Bank of New York, 8(1), 139-158. Retrieved from https://search.proquest.com/trade-journals/monetary-policy-transmission-residential/docview/210396886/se-2?accountid=13031

Meen, G. (1990), The Removal of Mortgage Market Constraints and the Implications for Econometric Modelling of UK House Prices, Oxford Bulletin of Economics and Statistics, 52, issue 1, p. 1-23,

Meen, G. (2002). The Time-Series Behavior of House Prices: A Transatlantic Divide?, Journal of Housing Economics, 11, issue 1, p. 1-23, ://EconPapers.repec.org/RePEc:eee:jhouse:v:11:y:2002:i:1:p:1-23

Official Statistics of Finland (OSF): Building and dwelling production [e-publication]. ISSN=1798-9590. Helsinki: Statistics Finland [referred: 4.4.2021]. Access method: http://www.stat.fi/til/ras/meta_en.html

Official Statistics of Finland (OSF): Building cost index [e-publication]. ISSN=1799-5841. Helsinki: Statistics Finland [referred: 4.4.2021]. Access method: http://www.stat.fi/til/rki/meta_en.html

Official Statistics of Finland (OSF): Consumer price index [e-publication]. ISSN=1799-0254. Helsinki: Statistics Finland [referred: 4.4.2021]. Access method: http://www.stat.fi/til/khi/meta_en.html

Official Statistics of Finland (OSF): Dwellings and housing conditions [e-publication]. ISSN=1798-6761. Overview 2017. Helsinki: Statistics Finland [referred: 16.3.2021]. Retrieved from: http://www.stat.fi/til/asas/2017/01/asas_2017_01_2018-10-10_tie_002_en.html

Official Statistics of Finland (OSF): Indebtedness [e-publication]. ISSN=2489-3285. Helsinki: Statistics Finland [referred: 3.4.2021]. Access method: http://www.stat.fi/til/velk/index_en.html

Official Statistics of Finland (OSF): Migration [e-publication]. ISSN=1797-6782. Helsinki: Statistics Finland [referred: 16.3.2021]. Retrieved from https://www.stat.fi/til/muutl/2019/02/muutl_2019_02_20201221_tie_001_en.html

Official Statistics of Finland (OSF): Population structure [e-publication]. ISSN=1797-5395. Helsinki: Statistics Finland [referred: 6.4.2021]. Access method: http://www.stat.fi/til/vaerak/meta_en.html

Official Statistics of Finland (OSF): Prices of dwellings in housing companies [e-publication]. ISSN=2323-8801. Helsinki: Statistics Finland [referred: 4.4.2021]. Access method: http://www.stat.fi/til/ashi/meta_en.html

Official Statistics of Finland (OSF): Quarterly national accounts [e-publication]. ISSN=1797-9765. Helsinki: Statistics Finland [referred: 4.4.2021]. Access method: http://www.stat.fi/til/ntp/meta_en.html

Oikarinen, E. (2005). Is housing overvalued in the Helsinki metropolitan. Helsinki: ETLA Discussion Papers, The Research Institute of the Finnish Economy (ETLA), No. 992.

Oikarinen, E. (2006). Price linkages between stock, bond and housing markets: Evidence from Finnish data (No. 1004). ETLA Discussion Papers.

Oikarinen, E. (2007). *Studies on housing price dynamics*. (Doctoral Thesis). Turku School of Economics. Series A-9:2007.

Oikarinen, E. (2009), Household borrowing and metropolitan housing price dynamics - Empirical evidence from Helsinki, Journal of Housing Economics, 18, issue 2, p. 126-139. https://EconPapers.repec.org/RePEc:eee:jhouse:v:18:y: 2009:i:2:p:126-139

Oikarinen, E. (2012) Empirical evidence on the reaction speeds of housing prices and sales to demand shocks, Journal of Housing Economics, Volume 21, Issue 1, Pages 41-54.

Oikarinen, E. (2014). Is urban land price adjustment more sluggish than housing price adjustment? Empirical evidence. Urban Studies, 51(8), 1686-1706.

Oikarinen, E. & Bourassa, S. C. & Hoesli, M. E. R. & Engblom, J. (2018). U.S. Metropolitan House Price Dynamics. Swiss Finance Institute Research Paper No. 17-72.

Poterba, J. M. (1984). Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach. Quarterly Journal of Economics, 99(4), 729–752. https://doi-org.pc124152.oulu.fi:9443/10.2307/1883123

Poterba, J.M., Weil, D., & Shiller, R. (1991). House Price Dynamics: The Role of Tax Policy and Demography. Brookings Papers on Economic Activity, 1991(2), 143-203. doi:10.2307/2534591

Pöörsisäätiö-Börsstiftelsen (2020) Sijoittajan vero-opas 2020. Retrieved from https://www.porssisaatio.fi/blog/books/sijoittajan-vero-opas/

Rayburn, W., Devaney, M. & Evans, R. (1987). A Test of Weak-Form Efficiency in Residential Real Estate Returns, Real Estate Economics, 15, issue 3, p. 220-233, https://EconPapers.repec.org/RePEc:bla:reesec:v:15:y:1987:i:3:p:220-233.

Riddel, M. (2004): "Housing-market disequilibrium: an examination of housing-market price and stock dynamics 1967-1998," Journal of Housing Economics, 13(2), 120–135.

SAS/ETS® (2014) 13.2 User's Guide "*The VARMAX Procedure*" SAS Institute Inc.

Smith, L. (1969). A Model of the Canadian Housing and Mortgage Markets. Journal of Political Economy, 77(5), 795-816. Retrieved March 11, 2021, from http://www.jstor.org/stable/1829968

Stiglitz, J. (1990). Symposium on Bubbles. The Journal of Economic Perspectives, 4(2), 13-18. Retrieved February 1, 2021, from http://www.jstor.org/stable/1942887

Suomen Pankki (n.d.). Talletusten_ja_lainojen_korot https://www.suomenpankki.fi /en/Statistics/mfi-balance-sheet/tables/rati-taulukoten/talletusten_ja_lainojen_ko rot_en/

Suomen Pankki (n.d.). Lainat_kotitalouksille https://www.suomenpankki.fi/en/Stat istics/ mfi-balance-sheet/tables/rati-taulukot-en/lainat_kotitalouksille_en/

Taipalus, K. (2006). A global house price bubble? Evaluation based on a new rent-price approach (No. 29/2006). Bank of Finland Research Discussion Papers.

Takala, K. & Pere, P. (1991). Testing the cointegration of house and stock prices in Finland, Finnish Economic Papers, 4, issue 1, p. 33-51

Topel, R., & Rosen, S. (1988). Housing Investment in the United States. Journal of Political Economy, 96(4), 718-740. Retrieved January 21, 2021, from http://www.jstor.org/stable/1830471

Vantaan Kaupungin Tilinpäätos (2009). Retrieved from: https://docplayer.fi/2038050-Vantaan-kaupungin-tilinpaatos-2009.html

Verohallinto (2015) Verotus 1865-2015 – näkökulmia suomalaiseen verotukseen. Retrieved from: https://www.vero.fi/globalassets/tietoa-verohallinnosta/esitys--ja-opetusmateriaalit/verotus_1865-2015.pdf

Verohallinto (n.d.). Kiinteistöveroprosentit https://www.vero.fi/henkiloasiakkaat/asuminen/kiinteistovero/nain-vero-muodostuu/kiinteistoveroprosenti/

Wang, S., & Tumbarello, P. (2010). What Drives House Prices in Australia? A+ L4584 Cross-Country Approach (Vol. 2010, No. 291, pp. 1-24). International Monetary Fund. Williams, Joseph T, (1991), Real Estate Development as an Option, The Journal of Real Estate Finance and Economics, 4, issue 2, p. 191-208, https://EconPapers.repec.org/RePEc:kap:jrefec:v:4:y:1991:i:2:p:191-208

Zhou, Z. (1997). Forecasting sales and price for existing single-family homes: A VAR model with error correction. The Journal of Real Estate Research, 14(1), 155-167. Retrieved from https://www.proquest.com/scholarly-journals/forecasting-sales-price-existing-single-family/docview/200268858/se-2?accountid=13031

APPENDICES

Appendix 1: Summary Statistics Capital Region of Helsinki

Table 1. Summary Statistics Capital Region of Helsinki

Explanatory variable	P	Y	UC	L	H	CC
Mean	4.393	6.911	0.028	0.313	7.829	4.915
Median	4.508	6.971	0.026	0.423	7.835	4.941
Maximum	4.669	7.259	0.066	0.830	8.426	4.974
Minimum	3.718	6.397	0.009	-0.448	7.017	4.839
Std. Dev.	0.270	0.227	0.015	0.430	0.238	0.044
Skewness	-1.089	-0.560	0.580	-0.326	-0.673	-0.284
Kurtosis	3.117	2.478	2.486	1.474	4.011	1.343
Jarque-Bera	19.812	6.357	6.698	11.473	11.814	12.784
Probability	0.000	0.042	0.035	0.003	0.003	0.002
Sum	439.342	691.127	2.780	31.340	782.855	491.482
Sum Sq. Dev.	7.211	5.084	0.023	18.266	5.609	0.195
Observations	100	100	100	100	100	100

P = real house prices, Y = real aggregate income, UC = real user cost, L= real loan to GDP ratio, H= housing stock, CC = real construction costs.

Appendix 2: Figures with the evolution of house price determinants 1995Q1-2019Q4 in differences

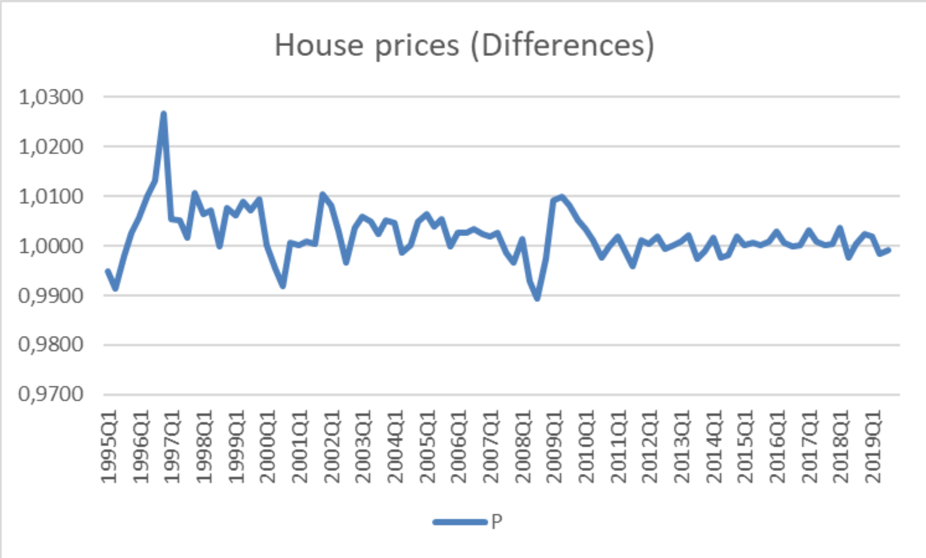

Figure 13. Evolution of house prices (differences) in CRH 1995Q1-2019Q4. (Source Tilastokeskus)

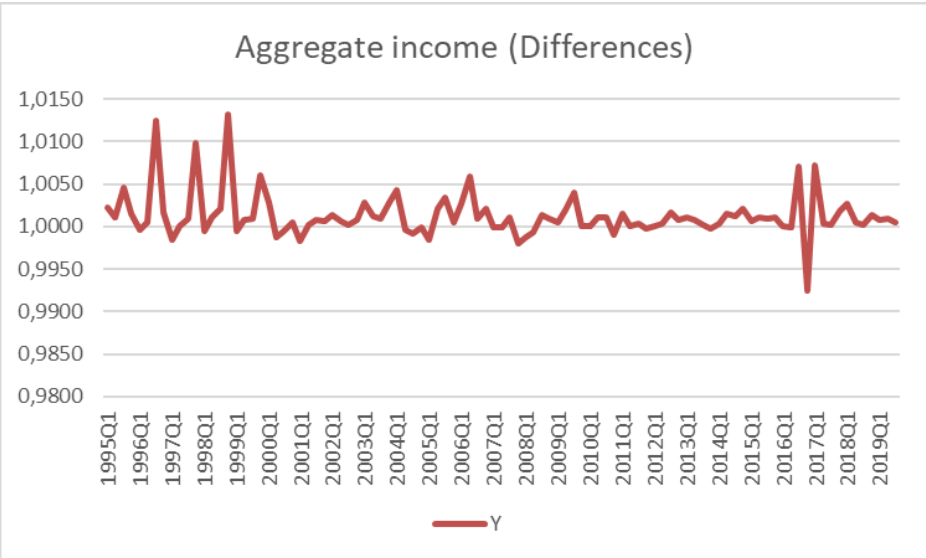

Figure 14. Evolution of aggregate income (differences) in CRH 1995Q1-2019Q4. (Source Tilastokeskus and professor Oikarinen)

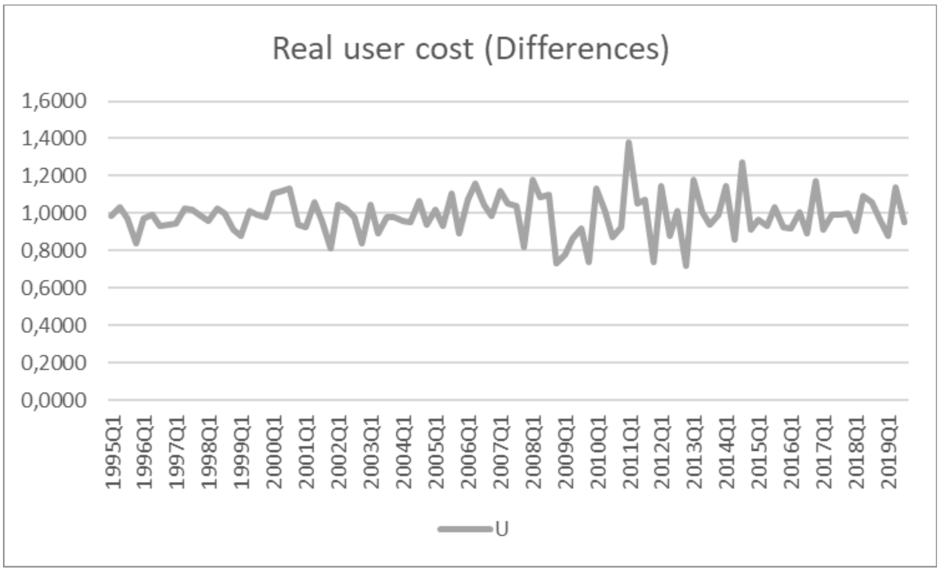

Figure 15. Evolution of real user cost (differences) in CRH 1995Q1-2019Q4. (Different sources)

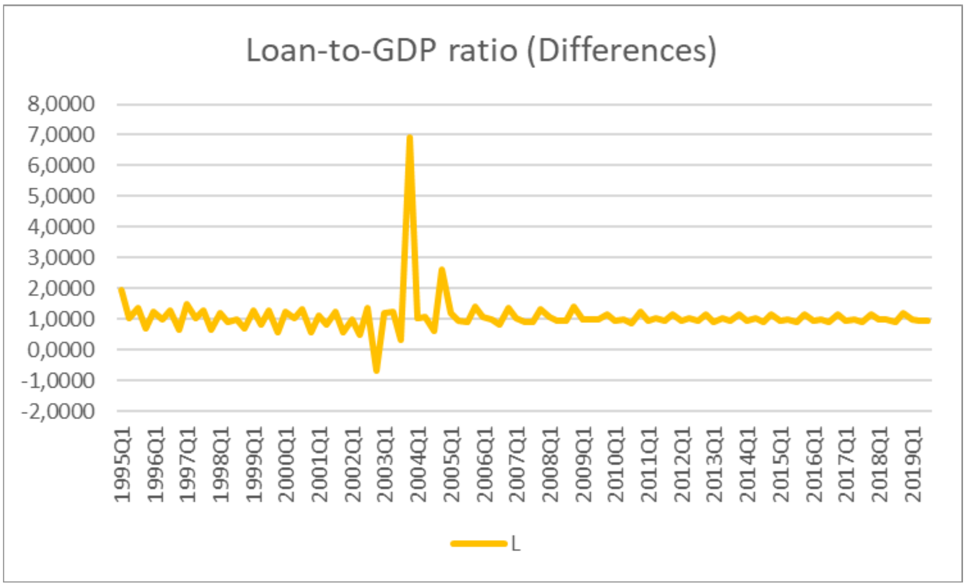

Figure 16. Evolution of loan-to-GDP ratio (differences) in Finland 1995Q1-2019Q4. (Sources: Tilastokeskus and Suomen Pankki)

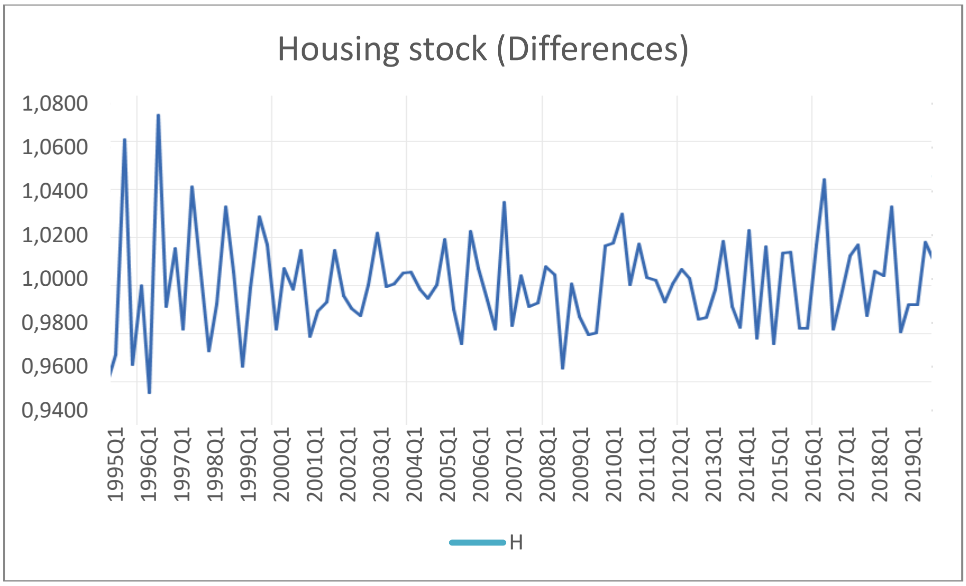

Figure 17. Evolution of housing stock (differences) in Finland (unadjusted) 1995Q1-2019Q4. (Source Tilastokeskus)

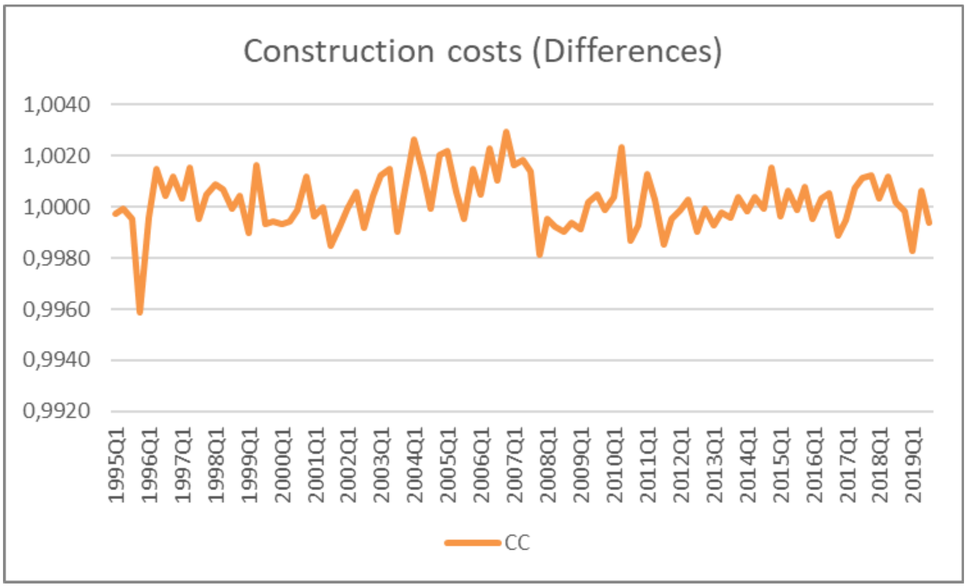

Figure 18. Evolution of construction costs (differences) in Finland 1995Q1-2019Q4. (Source Tilastokeskus)

Appendix 3: Lag order criteria

Table A3.1. Optimal lag order criteria: Model 1 (ln P= ln Y)

Lag	AIC	SBC	HQ
0	-3.307	-3.087	-3.218
1	-10.647	-10.318	-10.514
2	-10.870	-10.431*	-10.693*
3	-10.809	-10.261	-10.587
4	-10.844	-10.186	-10.578
5	-10.874	-10.107	-10.565
6	-10.888*	-10.011	-10.534
7	-10.814	-9.827	-10.415
8	-10.828	-9.732	-10.386

* indicates the lag order chosen by each criterion.

Table A3.2. Optimal lag order criteria: Model 2 (ln P= ln Y + ln L)

Lag	AIC	SBC	HQ
0	-4.937	-4.608	-4.805
1	-14.972	-14.397*	-14.740
2	-15.219	-14.396	-14.887*
3	-15.245	-14.176	-14.814
4	-15.302	-13.987	-14.771
5	-15.505*	-13.942	-14.874
6	-15.408	-13.599	-14.678
7	-15.394	-13.338	-14.564
8	-15.341	-13.038	-14.412

* indicates the lag order chosen by each criterion.

Table A3.3. Optimal lag order criteria: Model 3 (ln P= ln Y + ln L + ln U)

Lag	AIC	SBC	HQ
0	-12.605	-12.166	-12.428
1	-24.248	-23.371*	-23.894
2	-24.596	-23.281	-24.065*
3	-24.538	-22.784	-23.830
4	-24.571	-22.378	-23.686
5	-24.697*	-22.066	-23.635
6	-24.487	-21.417	-23.248
7	-24.445	-20.937	-23.029
8	-24.380	-20.433	-22.787

* indicates the lag order chosen by each criterion.

Appendix 4: VECM Model 1

Table A4.1. Summary of VECM. Sample (adjusted) 1995Q4-2019Q4

Long-term relationship	P	Y	C
β	1.000	-0.688	0.354
Standard. Errors		(0.139)	
t-statistics		[-4.954]	
Short-term relationship: Error Correction model		D(P)	D(Y)
Coint Eq. 1		-0.054	-0.034
		(0.014)	(0.016)
		[-3.956]	[-2.143]
P(-1)		0.517	0.149
		(0.106)	(0.125)
		[4.863]	[1.195]
P(-2)		-0.010	0.079
		(0.100)	(0.118)
		[-0.094]	[0.674]
Y(-1)		-0.047	-0.310
		(0.096)	(0.113)
		[-0.493]	[-2.746]
Y(-2)		-0.074	-0.185
		(0.095)	(0.112)
		[-0.779]	[-1.659]
C		0.016	0.014
		(0.003)	(0.004)
		[4.573]	[3.547]
S2		-0.007	-0.005
		(0.005)	(0.006)
		[-1.439]	[-0.822]
S3		-0.014	-0.008
		(0.004)	(0.005)
		[-2.997]	[-1.488]
S4		-0.018	-0.001
		(0.004)	(0.005)
		[-4.011]	[-0.145]
R-squared		0.492	0.147
Adj. R-squared		0.446	0.070
Sum sq. resids		0.020	0.027
S.E. equation		0.015	0.017
F-statistic		10.664	1.903
Log likelihood		274.998	259.454
Akaike AIC		-5.484	-5.164
Schwarz SC		-5.246	-4.925
Mean dependent		0.010	0.009
S.D. dependent		0.020	0.018
Determinant resid covariance (dof adj.)			6.04E-08
Determinant resid covariance			4.97E-08
Log likelihood			540.3865
Akaike information criterion			-10.72962
Schwarz criterion			-10.19875
Number of coefficients			20

Sample (adjusted): 1995Q4 2019Q4. Included observations: 97 after adjustments. Standard errors in ( ) & t-statistics in [ ].

Table A4.2. Dependent Variable: D(P)

D(P) = C(1)*(P(-1) - 0.687711276485*Y(-1) + 0.354108185137) + C(2)*D(P(-1)) + C(3)*D(P(-2)) + C(4)*D(Y(-1)) + C(5)*D(Y(-2)) + C(6) + C(7)*S2 + C(8)*S3 + C(9)*S4

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.054	0.014	-3.956	0.000
C(2)	0.517	0.106	4.863	0.000
C(3)	-0.010	0.100	-0.095	0.925
C(4)	-0.047	0.096	-0.493	0.623
C(5)	-0.074	0.095	-0.779	0.438
C(6)	0.016	0.003	4.573	0.000
C(7)	-0.007	0.005	-1.439	0.154
C(8)	-0.014	0.005	-2.997	0.004
C(9)	-0.018	0.004	-4.011	0.000

R-squared	0.492	Mean dependent var	0.010
Adjusted R-squared	0.446	S.D. dependent var	0.020
S.E. of regression	0.015	Akaike info criterion	-5.484
Sum squared residuals	0.020	Schwarz criterion	-5.246
Log likelihood	274.998	Hannan-Quinn criter.	-5.388
F-statistic	10.663	Durbin-Watson stat	2.041
Prob(F-statistic)	0.000		

Method: Least Squares (Gauss-Newton / Marquardt steps). Included observations: 97 after adjustments. C(1)= speed of adjustment coefficient. C(2) and C(3) are the one period and two period lagged house price. C(4) and C(5) are the lagged values for real aggregate income. C(6) is the constant. c(7), C(8) and C(9) are the seasonal dummies.

Table A4.3.1 Wald Test (Coeff. Diagnostics)

Test Statistic	Value	df	Probability
F-statistic	0.348	(2, 88)	0.707
Chi-square	0.696	2	0.706

Null Hypothesis: C(4)=C(5)=0

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4)	-0.047	0.096
C(5)	-0.074	0.095

Restrictions are linear in coefficients. C(4) and C(5) are the lagged values for real aggregate income.

Table A4.3.2 Wald Test (Coeff. Diagnostics)

Test Statistic	Value	df	Probability
F-statistic	6.006	(3, 88)	0.001
Chi-square	18.018	3	0.000
Null Hypothesis: C(7)=C(8)=C(9)=0			
Null Hypothesis Summary:			
Normalized Restriction (= 0)		Value	Std. Err.
C(7)		-0.007	0.005
C(8)		-0.014	0.005
C(9)		-0.018	0.004

Restrictions are linear in coefficients. C(7), C(8) and C(9) are the seasonal dummies.

Table A4.4. VEC Residual Serial Correlation LM Tests

Null hypothesis: No serial correlation at lag h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	4.779	4	0.311	1.205	(4, 170.0)	0.311
2	5.231	4	0.264	1.320	(4, 170.0)	0.264
Null hypothesis: No serial correlation at lags 1 to h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	4.779	4	0.311	1.205	(4, 170.0)	0.311
2	9.043	8	0.339	1.141	(8, 166.0)	0.339

*Edgeworth expansion corrected likelihood ratio statistic. Sample 1995Q1 2019Q4. Test includes 97 observations.

Table A4.5. VEC Residual Normality Tests

Orthogonalization: Cholesky (Lütkepohl)

Null Hypothesis: Residuals are multivariate normal

Component	Skewness	Chi-sq	df	Prob.*
1	0.040	0.027	1	0.870
2	1.152	21.456	1	0.000
Joint		21.482	2	0.000
Component	Kurtosis	Chi-sq	df	Prob.
1	4.583	10.135	1	0.002
2	7.034	65.779	1	0.000
Joint		75.914	2	0.000
Component	Jarque-Bera	df	Prob.	
1	10.162	2	0.006	
2	87.235	2	0.000	
Joint	97.396	4	0.000	

*Approximate p-values do not account for coefficient estimation. Sample: 1995Q1 2019Q4. Includes 97 observations.

Table A4.6. VEC Residual Heteroskedasticity Tests (Levels and Squares)

Joint test		Chi-sq	df	Prob.	
		72.839	39	0.001	
Individual components					
Dependent	R-squared	F(10,86)	Prob.	Chi-sq(10)	Prob.
res1*res1	0.413	4.488	0.000	40.041	0.000
res2*res2	0.138	1.018	0.442	13.339	0.422
res2*res1	0.246	2.086	0.023	23.886	0.032

Sample: 1995Q1 2019Q4. 97 observations are included.

Table A4.7. Granger-causality test

Null hypothesis	Obs	F-Statistic	Prob.
Y does not Granger Cause P	98	0.950	0.390
P does not Granger Cause Y		0.297	0.744

Test with 2 lags.

Appendix 5: VECM Model 2

Table A5.1. Summary of VECM. Sample (adjusted) 1995Q4-2019Q4

Long-term relationship	P	Y	L	C
β	1.000	-0.697	-0.237	0.493
Standard. Errors		(0.156)	(0.081)	
t-statistics		[-4.477]	[-2.941]	
Short-term relationship: Error Correction model		D(P)	D(Y)	D(L)
Coint Eq. 1		-0.045	0.000	0.169
		(0.023)	(0.000)	(0.042)
		[-1.968]	[NA]	[4.036]
P (-1)		0.617	0.193	-0.177
		(0.088)	(0.100)	(0.152)
		[6.997]	[1.932]	[-1.165]
Y(-1)		-0.033	-0.255	-0.076
		(0.097)	(0.110)	(0.167)
		[-0.340]	[-2.307]	[-0.453]
L(-1)		0.006	0.032	-0.154
		(0.057)	(0.065)	(0.098)
		[0.110]	[0.502]	[-1.576]
C		0.016	0.015	0.109
		(0.005)	(0.005)	(0.008)
		[3.438]	[2.872]	[13.450]
S2		-0.012	-0.014	-0.109
		(0.011)	(0.012)	(0.019)
		[-1.051]	[-1.092]	[-5.806]
S3		-0.018	-0.010	-0.110
		(0.005)	(0.006)	(0.009)
		[-3.629]	[-1.762]	[-12.875]
S4		-0.019	-0.003	-0.162
		(0.006)	(0.006)	(0.010)
		[-3.491]	[-0.413]	[-16.924]
R-squared		0.451	0.099	0.867
Adj. R-squared		0.408	0.029	0.857
Sum sq. resids		0.022	0.029	0.065
S.E. equation		0.016	0.018	0.027
F-statistic		10.552	1.408	83.970
Log likelihood		272.301	259.903	219.250
Akaike AIC		-5.394	-5.141	-4.311
Schwarz SC		-5.183	-4.930	-4.100
Mean dependent		0.009	0.009	0.009
S.D. dependent		0.020	0.018	0.071
Determinant resid covariance (dof adj.)			4.30E-11	
Determinant resid covariance			3.33E-11	
Log likelihood			764.217	
Akaike information criterion			-15.045	
Schwarz criterion			-14.333	
Number of coefficients			27	

Sample (adjusted): 1995Q4 2019Q4. Included observations: 98 after adjustments. Standard errors in ( ) & t-statistics in [ ]. Weakly exogeneity of aggregate income: the convergence is achieved after 5 iterations. LR test for binding restrictions (rank=1) with Chi-Square 2.784 and probability 0.095.

Table A5.2.1. Dependent Variable: D(P)

D(P) = C(1)*(P(-1) - 0.696646615686*Y(-1) - 0.237443013901*L(-1) + 0.493497766754) + C(2)*D(P(-1)) + C(3)*D(Y(-1)) + C(4)*D(L(-1)) + C(5) + C(6)*S2 + C(7)*S3 + C(8)*S4

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.056	0.024	-2.271	0.026
C(2)	0.617	0.088	6.997	0.000
C(3)	-0.033	0.097	-0.340	0.735
C(4)	0.006	0.057	0.110	0.913
C(5)	0.016	0.005	3.438	0.001
C(6)	-0.012	0.011	-1.051	0.296
C(7)	-0.018	0.005	-3.629	0.001
C(8)	-0.019	0.006	-3.491	0.001

R-squared	0.451	Mean dependent var	0.009
Adjusted R-squared	0.408	S.D. dependent var	0.020
S.E. of regression	0.016	Akaike info criterion	-5.394
Sum squared residuals	0.022	Schwarz criterion	-5.183
Log likelihood	272.301	Hannan-Quinn criter.	-5.309
F-statistic	10.552	Durbin-Watson stat	2.083
Prob(F-statistic)	0.000		

Method: Least Squares (Gauss-Newton / Marquardt steps). Included observations: 98 after adjustments. C(1)= speed of adjustment coefficient. C(2) is the one period lagged house price. C(3) are the lagged value for real aggregate income. C(4) is the lagged value for Loan-to-GDP ratio. C(5) is the constant. C(6), C(7) and C(8) are seasonal dummies.

Table A5.2.2. Wald Test (Coeff. Diagnostics)

Test Statistic	Value	df	Probability
F-statistic	6.999	(3, 90)	0.000
Chi-square	20.999	3	0.000

Null Hypothesis: C(6)=C(7)=C(8)=0

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(6)	-0.018	0.005
C(7)	-0.019	0.006
C(8)	-0.011	0.011

Restrictions are linear in coefficients. C(6), C(7) and C(8) are the seasonal dummies.

Table A5.3. VEC Residual Serial Correlation LM Tests

Null hypothesis: No serial correlation at lag h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	8.027	9	0.531	0.894	(9, 207.0)	0.532
Null hypothesis: No serial correlation at lags 1 to h						
Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	8.027	9	0.531	0.894	(9, 207.0)	0.532

*Edgeworth expansion corrected likelihood ratio statistic. Sample 1995Q1 2019Q4. Test includes 98 observations.

Table A5.4. VEC Residual Normality Tests

Orthogonalization: Cholesky (Lütkepohl)
Null Hypothesis: Residuals are multivariate normal

Component	Skewness	Chi-sq	df	Prob.*
1	0.090	0.134	1	0.715
2	1.292	27.244	1	0.000
3	0.329	1.772	1	0.183
Joint		29.150	3	0.000
Component	Kurtosis	Chi-sq	df	Prob.
1	4.560	9.940	1	0.002
2	7.032	66.396	1	0.000
3	3.871	3.095	1	0.079
Joint		79.431	3	0.000
Component	Jarque-Bera	df	Prob.	
1	10.074	2	0.007	
2	93.640	2	0.000	
3	4.867	2	0.088	
Joint	108.582	6	0.000	

*Approximate p-values do not account for coefficient estimation. Sample: 1995Q1 2019Q4. Includes 98 observations.

Table A5.5. VEC Residual Heteroskedasticity Tests (Levels and Squares)

Joint test:		Chi-sq	df	Prob.	
		124.217	66	0.000	
Individual components:					
Dependent	R-squared	F(11,86)	Prob.	Chi-sq(10)	Prob.
res1*res1	0.545	9.372	0.000	53.429	0.000
res2*res2	0.105	0.921	0.525	10.324	0.502
res3*res3	0.214	2.126	0.027	20.955	0.034
res2*res1	0.302	3.386	0.001	29.618	0.002
res3*res1	0.468	6.883	0.000	45.884	0.000
res3*res2	0.306	3.454	0.001	30.027	0.002

Sample: 1995Q1 2019Q4. 98 observations are included.

Table A5.6. Granger-causality test

Null hypothesis	Obs	F-Statistic	Prob.
Y does not Granger Cause P	99	0.186	0.668
P does not Granger Cause Y		0.089	0.766
L does not Granger Cause P	99	0.634	0.428
P does not Granger Cause L		9.219	0.003

Test with 1 lag.

Appendix 6: VECM Model 3

Table A6.1. Summary of VECM. Sample (adjusted) 1995Q4-2019Q4

Long-term relationship	P	Y	L	U	C
β	1.000	-0.492	-0.302	0.000	-0.905
Standard. Errors		(0.113)	(0.057)		
t-statistics		[-4.355]	[-5.293]		
Short-term relationship: Error Correction model		D(P)	D(Y)	D(L)	D(U)
Coint Eq.1		-0.068	-0.093	0.175	0.017
		(0.025)	(0.031)	(0.048)	(0.004)
		[-2.761]	[-3.017]	[3.618]	[4.003]
P (-1)		0.440	0.233	-0.235	-0.025
		(0.116)	(0.144)	(0.227)	(0.020)
		[3.787]	[1.616]	[-1.035]	[-1.274]
P(-2)		0.050	0.138	-0.015	0.027
		(0.102)	(0.126)	(0.199)	(0.017)
		[0.488]	[1.093]	[-0.073]	[1.566]
Y(-1)		-0.008	-0.350	-0.082	-0.004
		(0.095)	(0.118)	(0.186)	(0.016)
		[-0.088]	[-2.967]	[-0.438]	[-0.239]
Y(-2)		-0.075	-0.167	-0.200	-0.002
		(0.090)	(0.112)	(0.176)	(0.015)
		[-0.837]	[-1.494]	[-1.135]	[-0.154]
L(-1)		0.001	0.075	-0.173	-0.028
		(0.055)	(0.069)	(0.108)	(0.009)
		[0.020]	[1.091]	[-1.598]	[-2.992]
L(-2)		0.135	0.147	-0.068	-0.016
		(0.056)	(0.070)	(0.111)	(0.010)
		[2.388]	[2.082]	[-0.616]	[-1.667]
U(-1)		-1.513	1.468	-2.511	-0.029
		(0.636)	(0.790)	(1.244)	(0.108)
		[-2.378]	[1.857]	[-2.018]	[-0.270]
U(-2)		-0.909	0.430	-1.221	-0.081
		(0.648)	(0.805)	(1.268)	(0.110)
		[-1.402]	[0.534]	[-0.963]	[-0.741]
C		0.0165	0.0186	0.111	-0.005
		(0.004)	(0.005)	(0.008)	(0.001)
		[3.838]	[3.475]	[13.131]	[-6.723]
S2		-0.005	-0.008	-0.116	0.008
		(0.011)	(0.013)	(0.021)	(0.002)
		[-0.500]	[-0.598]	[-5.596]	[4.550]
S3		-0.033	-0.029	-0.104	0.006
		(0.008)	(0.010)	(0.016)	(0.001)
		[-3.966]	[-2.756]	[-6.305]	[4.285]
S4		-0.015	-0.003	-0.162	0.005
		(0.005)	(0.007)	(0.010)	(0.001)
		[-2.77]	[-0.466]	[-15.606]	[5.159]
R-squared		0.574	0.200	0.872	0.465
Adj. R-squared		0.514	0.086	0.854	0.388
Sum sq. resids		0.016	0.025	0.063	0.000
S.E. equation		0.014	0.017	0.027	0.002
F-statistic		9.449	1.750	47.855	6.079
Log likelihood		283.564	262.534	218.482	455.708
Akaike AIC		-5.579	-5.145	-4.237	-9.128
Schwarz SC		-5.234	-4.800	-3.892	-8.783
Mean dependent		0.010	0.009	0.009	-0.001
S.D. dependent		0.020	0.018	0.072	0.003
Determinant resid covariance (dof adj.)			1.75E-16		

Determinant resid covariance	9.83E-17
Log likelihood	1237.069
Akaike information criterion	-24.352
Schwarz criterion	-22.866
Number of coefficients	56

Sample (adjusted): 1995Q3 2019Q4. Included observations: 97 after adjustments. Standard errors in ( ) & t-statistics in [ ]. Convergence achieved after 6 iterations. LR test for binding restrictions (rank = 1): Chi-square(1) is 2.016 with probability 0.156.

Table A6.2. Dependent Variable: D(P)

D(P) = C(1)*(P(-1) - 0.492*Y(-1) - 0.302*L(-1) - 0.905) + C(2)*D(P(-1)) + C(3)*D(P(-2)) + C(4)*D(Y(-1)) + C(5)*D(Y(-2)) + C(6)*D(L(-1)) + C(7)*D(L(-2)) + C(8)*D(U(-1)) + C(9)*D(U(-2)) + C(10) + C(11)*S2 + C(12)*S3 + C(13)*S4

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.068	0.025	-2.761	0.007
C(2)	0.440	0.116	3.787	0.000
C(3)	0.050	0.102	0.488	0.627
C(4)	-0.008	0.095	-0.088	0.930
C(5)	-0.075	0.090	-0.837	0.405
C(6)	0.001	0.055	0.020	0.984
C(7)	0.135	0.057	2.388	0.019
C(8)	-1.513	0.636	-2.378	0.020
C(9)	-0.909	0.648	-1.402	0.165
C(10)	0.017	0.004	3.838	0.000
C(11)	-0.005	0.011	-0.500	0.618
C(12)	-0.033	0.008	-3.966	0.000
C(13)	-0.015	0.005	-2.773	0.007
R-squared	0.574	Mean dependent var		0.010
Adjusted R-squared	0.514	S.D. dependent var		0.020
S.E. of regression	0.014	Akaike info criterion		-5.579
Sum squared residuals	0.016	Schwarz criterion		-5.234
Log likelihood	283.564	Hannan-Quinn criter.		-5.440
F-statistic	9.450	Durbin-Watson stat		1.927
Prob(F-statistic)	0.000			

Method: Least Squares (Gauss-Newton / Marquardt steps). Included observations: 98 after adjustments. C(1)= speed of adjustment coefficient. C(2) and C(3) is the lagged values for house price. C(4) and C(5) are the lagged value for real aggregate income. C(6) and C(7) are the lagged value for Loan-to-GDP ratio. C(8) and C(9) are the lagged values for real user cost. C(10) is the constant. And C(11), C(12) and C(13) are the seasonal dummies, S2, S3 and S4.

Table A6.2.1. Wald Test (Coeff. Diagnostics)

Test Statistic	Value	df	Probability
F-statistic	0.197	(2, 91)	0.822
Chi-square	0.393	2	0.822

Null Hypothesis: C(6)=C(7)=0

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(6)	0.044	0.101
C(7)	-0.032	0.100

Restrictions are linear in coefficients. C(6) and C(7) are the lagged values for the loan-to-GDP ratio.

Table A6.3. VEC Residual Serial Correlation LM Tests

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	20.239	16	0.210	1.281	(16, 235.9)	0.210
2	23.846	16	0.093	1.521	(16, 235.9)	0.093

Null hypothesis: No serial correlation at lags 1 to h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	20.239	16	0.210	1.281	(16, 235.9)	0.210
2	49.633	32	0.024	1.605	(32, 270.8)	0.024

*Edgeworth expansion corrected likelihood ratio statistic. Sample 1995Q1 2019Q4. Test includes 97 observations.

Table A6.4. VEC Residual Normality Tests

Orthogonalization: Cholesky (Lütkepohl)

Null Hypothesis: Residuals are multivariate normal

Component	Skewness	Chi-sq	df	Prob.*
1	0.192	0.595	1	0.440
2	0.870	12.229	1	0.001
3	0.442	3.154	1	0.076
4	-0.268	1.157	1	0.282
Joint		17.135	4	0.002
Component	Kurtosis	Chi-sq	df	Prob.
1	3.843	2.872	1	0.090
2	5.850	32.818	1	0.000
3	4.043	4.393	1	0.036
4	3.266	0.285	1	0.593
Joint		40.368	4	0.000
Component	Jarque-Bera	df	Prob.	
1	3.468	2	0.177	
2	45.046	2	0.000	
3	7.547	2	0.023	
4	1.442	2	0.486	
Joint	57.503	8	0.000	

*approximate p-values do not account for coefficient estimation. Sample: 1995Q1 2019Q4. Includes 97 observations.

Table A6.5. VEC Residual Heteroskedasticity Tests (Levels and Squares)

Joint test:		Chi-sq	df	Prob.	
		266.939	210	0.005	
Individual components:					
Dependent	R-squared	F(21,75)	Prob.	Chi-sq(10)	Prob.
res1*res1	0.529	4.010	0.000	51.307	0.000
res2*res2	0.186	0.817	0.691	18.059	0.645
res3*res3	0.150	0.631	0.882	14.561	0.844
res4*res4	0.479	3.282	0.000	46.451	0.001
res2*res1	0.292	1.476	0.112	28.370	0.130
res3*res1	0.371	2.105	0.010	35.973	0.022
res3*res2	0.300	1.532	0.092	29.114	0.111
res4*res1	0.299	1.526	0.094	29.036	0.113
res4*res2	0.178	0.772	0.743	17.245	0.696
res4*res3	0.370	2.100	0.010	35.921	0.022

Sample: 1995Q1 2019Q4. 97 observations are included.

Table A6.6. Granger-causality test

Null hypothesis	Obs	F-Statistic	Prob.
Y does not Granger Cause P	98	0.950	0.390
P does not Granger Cause Y		0.297	0.744
L does not Granger Cause P	98	0.151	0.860
P does not Granger Cause L		5.284	0.007
U does not Granger Cause P	98	8.083	0.001
P does not Granger Cause U		0.323	0.725

Test with 2 lags.